10. Math 235 Fall 2023, Worksheet 10: Applications of linear algebra

This worksheet is devoted to linear algebra, and specifically to its uses in contest-style problems (often of combinatorial or arithmetical nature).

This is **not** an introduction to linear algebra, as there are several good textbooks for that: Strickland's notes [Strick21] are a well-written and rigorous introduction to matrices (despite their applied audience), whereas the textbook [LaNaSc16] exposes the subject from the point of view of vector spaces. Axler's book [Axler23] is also recommended, even though it bungles the definition of a polynomial[1]. Other (more advanced) texts are [Treil21], [Griffin20], [Camero08], [StoLui18] and [Kuttle22].[2]

Thus, I assume that you are familiar with the basics of linear algebra (although I will insert the occasional reminder), and set out to show how it can be used in some unexpected places.

As before, $\mathbb{N}$ means the set $\{0,1,2,\ldots\}$.

10.1. A note on fields

One of the main concepts in abstract algebra is that of a *field*. Roughly speaking, a field is a set of "numbers" (in a sufficiently wide sense), which can be added, subtracted, multiplied and divided (except by zero). These four operations are assumed to satisfy certain axioms (commutativity, associativity, distributivity, existence of 0 and 1, and of course we want subtraction to undo addition and division to undo multiplication). See any textbook on abstract algebra for details (e.g., [Aluffi21, Chapter 3] or [Steinb06, Chapter 7 onwards] or [Grinbe23a, §2.5 onwards]). The sets $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are fields, with the usual four operations. The set $\mathbb{Z}$ is not, since integers cannot always be divided (at least not without leaving $\mathbb{Z}$ or incurring remainders). Other fields exist, and we will meet one more on this very worksheet; but the most important ones are $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$.

The main concepts of linear algebra (matrices, vector spaces, linear maps, determinants) can be defined over any field: If $\mathbb{F}$ is any field (e.g., one of $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$), then we can define

- matrices over $\mathbb{F}$ (these are matrices whose entries belong to $\mathbb{F}$);
- vector spaces over $\mathbb{F}$ (these are vector spaces in which the vectors can be scaled by the elements of $\mathbb{F}$);
- linear maps (these are maps $f : V \to W$ that satisfy $f(v_1 + v_2) = f(v_1) + f(v_2)$ for all $v_1, v_2 \in V$ and $f(\lambda v) = \lambda f(v)$ for all $\lambda \in \mathbb{F}$ and $v \in V$);

[1] Every time Axler says "polynomial", read "polynomial function" instead.

[2] This list restricts itself to the sources openly available online and written in English; otherwise it would go on for several more lines.

- and so on.

When $\mathbb{F} = \mathbb{R}$, these concepts become the usual "real" concepts (i.e., matrices with real entries, real vector spaces, real linear maps, etc.) known from basic courses on linear algebra. Likewise, for $\mathbb{F} = \mathbb{C}$, we obtain complex matrices, complex vector spaces, etc.[3]. In general, when doing linear algebra using a field $\mathbb{F}$, one commonly says that one is "working over $\mathbb{F}$", and one refers to the elements of $\mathbb{F}$ as "*scalars*".

10.2. Linear dependence, bases, spanning

Some of the most fundamental facts in linear algebra concern lists of vectors in vector spaces. Such lists can be linearly dependent or independent; they have spans (which are subspaces); they can be a basis of the vector space. Furthermore, finite-dimensional vector spaces have bases and dimensions[4]. This all can be found in any good textbook on linear algebra (e.g., [LaNaSc16, Chapter 5] or [Axler23, Chapter 2]). Let me highlight one particular result for its usefulness:

Theorem 10.2.1. Let $k, n \in \mathbb{N}$. Assume you have a list of k vectors in an n-dimensional vector space V.

(a) If $k > n$, then your k vectors are linearly dependent.

(b) If $k = n$ and if your k vectors are linearly independent, then they form a basis of V.

(c) If $k < n$, then your k vectors cannot span V.

(d) If $k = n$ and if your k vectors span V, then they form a basis of V.

Proof. Part **(a)** is [Griffin20, Theorem 1.59]. Part **(b)** is [Griffin20, Theorem 1.63] or [LaNaSc16, Theorem 5.4.4 part 3]. Part **(c)** follows easily from [LaNaSc16, Theorem 5.2.9]. Part **(d)** is [LaNaSc16, Theorem 5.4.4 part 2]. □

The four parts of Theorem 10.2.1 are sometimes called the *pigeonhole principles for vector spaces*, due to their similarity to the pigeonhole principles for finite sets (Theorem 3.0.1 in Worksheet 3). Just like the latter, they require finiteness (k and n

[3] For example, the matrix $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ can be viewed both as a real matrix and as a complex matrix (and even as a rational matrix). As a real matrix, it cannot be diagonalized, since its eigenvalues $1+i$ and $1-i$ are not real. But as a complex matrix, it can be diagonalized:

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} 1-i & 0 \\ 0 & 1+i \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}^{-1}.$$

[4] Infinite-dimensional vector spaces have bases and dimensions, too, at least if you believe in the axiom of choice. But this is both harder to prove and far less useful than the finite case.

must be finite, although V will usually not be finite as a set), but finite-dimensional vector spaces are widespread in mathematics, so there is no shortage of situations to apply them to.

Part **(a)** alone is surprisingly useful. The following two "exercises" are combinatorial results with known names and a long history; both are rather hard to prove by elementary combinatorics. Using Theorem 10.2.1 **(a)**, we will prove them both with fairly little effort. The first is known as the *easy Lindström theorem* (the hard one is Exercise 10.7.1):

Exercise 10.2.1. Let $n \in \mathbb{N}$. Let S be an n-element set. Let $A_1, A_2, \ldots, A_{n+1}$ be $n+1$ nonempty subsets of S. Prove that there exist two disjoint nonempty subsets I and J of $\{1, 2, \ldots, n+1\}$ such that

$$\bigcup_{i \in I} A_i = \bigcup_{i \in J} A_i.$$

(Recall that $\bigcup_{i \in I} A_i$ denotes the union of the A_i for all $i \in I$. For example, if $I = \{2, 3, 6\}$, then $\bigcup_{i \in I} A_i = A_2 \cup A_3 \cup A_6$. Similarly, $\bigcup_{i \in J} A_i$ is understood.)

Example 10.2.2. For instance, let $n = 5$ and $S = \{1, 2, 3, 4, 5\}$. Let

$$\begin{aligned} A_1 &= \{1, 2, 3\}, & A_2 &= \{2, 3, 5\}, & A_3 &= \{2, 4\}, \\ A_4 &= \{1, 5\}, & A_5 &= \{2, 3\}, & A_6 &= \{1\}. \end{aligned}$$

Then, $A_5 \cup A_6 = A_1$. In other words, the two disjoint nonempty subsets $I = \{5, 6\}$ and $J = \{1\}$ of $\{1, 2, 3, 4, 5, 6\}$ satisfy $\bigcup_{i \in I} A_i = \bigcup_{i \in J} A_i$.

Example 10.2.3. For another example, let $n = 5$ and $S = \{1, 2, 3, 4, 5\}$. Let

$$\begin{aligned} A_1 &= \{1, 3, 4\}, & A_2 &= \{2, 3, 5\}, & A_3 &= \{2, 4\}, \\ A_4 &= \{1, 5\}, & A_5 &= \{2, 3\}, & A_6 &= \{1\}. \end{aligned}$$

Then, $A_1 \cup A_2 = A_3 \cup A_4 \cup A_5$. In other words, the two disjoint nonempty subsets $I = \{1, 2\}$ and $J = \{3, 4, 5\}$ of $\{1, 2, 3, 4, 5, 6\}$ satisfy $\bigcup_{i \in I} A_i = \bigcup_{i \in J} A_i$.

Example 10.2.4. We cannot allow the sets $A_1, A_2, \ldots, A_{n+1}$ in Exercise 10.2.1 to be empty. In fact, if we did so, then we could set $S = \{1, 2, \ldots, n\}$ and

$$\begin{aligned} A_1 &= \{1\}, & A_2 &= \{2\}, & &\ldots, & A_n &= \{n\}, \\ A_{n+1} &= \varnothing, \end{aligned}$$

and then there would not be any two disjoint nonempty subsets I and J of $\{1, 2, \ldots, n+1\}$ such that $\bigcup_{i \in I} A_i = \bigcup_{i \in J} A_i$.

Solution to Exercise 10.2.1. WLOG assume that $S = \{1, 2, \ldots, n\}$ (otherwise, just rename the elements).

We shall use vectors – specifically, row vectors of size n with real entries. Such vectors are written as $(a_1, a_2, \ldots, a_n)$, and the vector space consisting of these vectors is called $\mathbb{R}^n$. The j-th entry of such a vector is called its *j-th coordinate*.

For each subset J of S, we define the *indicator vector* $e_J \in \mathbb{R}^n$ as follows: For each $i \in S = \{1, 2, \ldots, n\}$, the i-th coordinate of e_J shall be

$$\begin{cases} 1, & \text{if } i \in J; \\ 0, & \text{if } i \notin J. \end{cases}$$

For example:

- If $n = 5$ and $J = \{1, 3, 4\}$, then $e_J = (1, 0, 1, 1, 0)$.
- If $n = 5$ and $J = \{2, 4\}$, then $e_J = (0, 1, 0, 1, 0)$.
- If $n = 5$ and $J = \varnothing$, then $e_J = (0, 0, 0, 0, 0)$.
- If $n = 5$ and $J = S = \{1, 2, 3, 4, 5\}$, then $e_J = (1, 1, 1, 1, 1)$.

This indicator vector e_J is also known as "incidence vector" or "characteristic vector", but we prefer the name "indicator vector" since it is the most informative (the entries of e_J literally indicate what numbers belong to J, like an indicator light on an electric appliance).

Thus, we have $n + 1$ indicator vectors $e_{A_1}, e_{A_2}, \ldots, e_{A_{n+1}}$ all belonging to the n-dimensional vector space $\mathbb{R}^n$ (of row vectors). By Theorem 10.2.1 **(a)** (applied to $m = n + 1$), we thus conclude that they are linearly dependent (since $n + 1 > n$). In other words, there exist real numbers $c_1, c_2, \ldots, c_{n+1}$, not all zero, such that

$$c_1 e_{A_1} + c_2 e_{A_2} + \cdots + c_{n+1} e_{A_{n+1}} = \mathbf{0} \tag{1}$$

(where $\mathbf{0}$ denotes the zero vector). Consider these numbers $c_1, c_2, \ldots, c_{n+1}$.

Let I and J be the two subsets of $\{1, 2, \ldots, n + 1\}$ defined by

$$I = \{i \in \{1, 2, \ldots, n+1\} \mid c_i > 0\} \qquad \text{and}$$
$$J = \{i \in \{1, 2, \ldots, n+1\} \mid c_i < 0\}.$$

These two subsets I and J are disjoint (since a single i cannot satisfy $c_i > 0$ and $c_i < 0$ at the same time). Moreover, we can easily rewrite the equality (1) as

$$\sum_{i \in I} c_i e_{A_i} = -\sum_{i \in J} c_i e_{A_i}, \tag{2}$$

since (1) leads to

$$\begin{aligned}
\mathbf{0} &= c_1 e_{A_1} + c_2 e_{A_2} + \cdots + c_{n+1} e_{A_{n+1}} = \sum_{i \in \{1,2,\ldots,n+1\}} c_i e_{A_i} \\
&= \underbrace{\sum_{\substack{i \in \{1,2,\ldots,n+1\}; \\ c_i > 0}} c_i e_{A_i}}_{\substack{= \sum_{i \in I} \\ \text{(by the definition of } I\text{)}}} + \sum_{\substack{i \in \{1,2,\ldots,n+1\}; \\ c_i = 0}} \underbrace{c_i}_{=0} e_{A_i} + \underbrace{\sum_{\substack{i \in \{1,2,\ldots,n+1\}; \\ c_i < 0}} c_i e_{A_i}}_{\substack{= \sum_{i \in J} \\ \text{(by the definition of } J\text{)}}} \\
&\qquad \left(\begin{array}{c} \text{since each } i \in \{1,2,\ldots,n+1\} \text{ satisfies} \\ \text{either } c_i > 0 \text{ or } c_i = 0 \text{ or } c_i < 0 \end{array} \right) \\
&= \sum_{i \in I} c_i e_{A_i} + \underbrace{\sum_{\substack{i \in \{1,2,\ldots,n+1\}; \\ c_i = 0}} 0 e_{A_i}}_{=0} + \sum_{i \in J} c_i e_{A_i} \\
&= \sum_{i \in I} c_i e_{A_i} + \sum_{i \in J} c_i e_{A_i}.
\end{aligned}$$

Now comes a crucial but simple observation: The equality (2) entails

$$\bigcup_{i \in I} A_i = \bigcup_{i \in J} A_i. \tag{3}$$

Proof of (3). Let $k \in S$. Thus, $k \in S = \{1, 2, \ldots, n\}$. Recall that the k-th coordinate of a given indicator vector e_{A_i} is

$$\begin{cases} 1, & \text{if } k \in A_i; \\ 0, & \text{if } k \notin A_i. \end{cases}$$

Thus, the k-th coordinate of the vector $\sum_{i \in I} c_i e_{A_i}$ equals

$$\sum_{i \in I} c_i \begin{cases} 1, & \text{if } k \in A_i; \\ 0, & \text{if } k \notin A_i \end{cases} = \sum_{\substack{i \in I; \\ k \in A_i}} \underbrace{c_i \cdot 1}_{=c_i} + \underbrace{\sum_{\substack{i \in I; \\ k \notin A_i}} c_i \cdot 0}_{=0} = \sum_{\substack{i \in I; \\ k \in A_i}} c_i.$$

This sum is 0 when $k \notin \bigcup_{i \in I} A_i$ (because in this case, the sum is empty[5]), and positive otherwise (because if $k \in \bigcup_{i \in I} A_i$, then the sum $\sum_{\substack{i \in I; \\ k \in A_i}} c_i$ is nonempty[6], and all its

[5] since $k \notin \bigcup_{i \in I} A_i$ means that there exists no $i \in I$ satisfying $k \in A_i$

[6] since $k \in \bigcup_{i \in I} A_i$ means that there exists some $i \in I$ satisfying $k \in A_i$

addends are positive[7]). Thus, we have shown that the k-th coordinate of the vector $\sum_{i\in I} c_i e_{A_i}$ is 0 when $k \notin \bigcup_{i\in I} A_i$ and positive otherwise. Hence, the k-th coordinate of the vector $\sum_{i\in I} c_i e_{A_i}$ is nonzero exactly when $k \in \bigcup_{i\in I} A_i$.

An analogous argument shows that the k-th coordinate of the vector $\sum_{i\in J} c_i e_{A_i}$ is nonzero exactly when $k \in \bigcup_{i\in J} A_i$. (The only difference is that now it will be negative, rather than positive, in the case when $k \in \bigcup_{i\in J} A_i$.)

However, the equality (2) shows that the two vectors $\sum_{i\in I} c_i e_{A_i}$ and $\sum_{i\in J} c_i e_{A_i}$ are equal up to a factor of -1. Thus, in particular, the k-th coordinate of the former vector equals the k-th coordinate of the latter vector times -1. Hence, the k-th coordinate of the former vector is nonzero if and only if the k-th coordinate of the latter vector is nonzero. In other words, we have $k \in \bigcup_{i\in I} A_i$ if and only if $k \in \bigcup_{i\in J} A_i$ (since the k-th coordinate of the vector $\sum_{i\in I} c_i e_{A_i}$ is nonzero exactly when $k \in \bigcup_{i\in I} A_i$, and since the k-th coordinate of the vector $\sum_{i\in J} c_i e_{A_i}$ is nonzero exactly when $k \in \bigcup_{i\in J} A_i$).

Forget that we fixed k. We thus have shown that for each $k \in S$, we have $k \in \bigcup_{i\in I} A_i$ if and only if $k \in \bigcup_{i\in J} A_i$. In other words, the sets $\bigcup_{i\in I} A_i$ and $\bigcup_{i\in J} A_i$ contain the exact same elements of S. Since both $\bigcup_{i\in I} A_i$ and $\bigcup_{i\in J} A_i$ are subsets of S, this entails that they are identical. In other words, $\bigcup_{i\in I} A_i = \bigcup_{i\in J} A_i$. This proves (3). □

We are now almost done: We have constructed two disjoint subsets I and J of $\{1, 2, \ldots, n+1\}$ such that

$$\bigcup_{i\in I} A_i = \bigcup_{i\in J} A_i.$$

It remains to prove that I and J are nonempty.

Here, we need to recall that the numbers $c_1, c_2, \ldots, c_{n+1}$ are not all zero. Thus, at least one of them is positive or negative. In other words, at least one of the sets I and J is nonempty. Let us WLOG assume that $J \neq \varnothing$ (since the case $I \neq \varnothing$ is analogous). We shall now show that $I \neq \varnothing$ as well. Indeed, from $J \neq \varnothing$, we see that $\bigcup_{i\in J} A_i$ is a nonempty union of nonempty sets (since all $n+1$ sets $A_1, A_2, \ldots, A_{n+1}$ are nonempty), and thus is nonempty itself. In other words, $\bigcup_{i\in I} A_i$ is nonempty (since $\bigcup_{i\in I} A_i = \bigcup_{i\in J} A_i$). But this is only possible if $I \neq \varnothing$ (since $I = \varnothing$ would lead

[7] since $c_i > 0$ for each $i \in I$

to $\bigcup_{i \in I} A_i = \bigcup_{i \in \varnothing} A_i = \varnothing$). Hence, we have shown that $I \neq \varnothing$. Thus, both I and J are nonempty, and this completes our solution to Exercise 10.2.1. $\square$

The next exercise is known as the the *nonuniform Fisher inequality* ([BabFra23, Theorem 4.1]):

Exercise 10.2.2. Let k, n and m be three positive integers. In a town with n inhabitants, there are m clubs, no two of which have the exact same set of members. Assume that any two distinct clubs share exactly k members. Prove that $m \leq n$.

Example 10.2.5. Let $n = 3$ and let the inhabitants be $1, 2, 3$. Consider the three clubs

$$\{1,2\},\ \{1,3\},\ \{2,3\}$$

(of course, we regard each club as the set of its members). Any two of them share exactly 1 member. Exercise 10.2.2 claims that we cannot find more than 3 clubs with this property.

Example 10.2.6. Let $n = 4$ and let the inhabitants be $1, 2, 3, 4$. Consider the four clubs

$$\{1,2,3\},\ \{2,3,4\},\ \{3,4,1\},\ \{4,1,2\}.$$

Any two of them share exactly 2 members. Exercise 10.2.2 claims that we cannot find more than 4 clubs with this property.

Solution idea to Exercise 10.2.2. Let the n inhabitants be called $1, 2, \ldots, n$. Thus, each club is a subset of $\{1, 2, \ldots, n\}$.

Let $C_1, C_2, \ldots, C_m$ be our m clubs (regarded as subsets of $\{1, 2, \ldots, n\}$). Let $v_1, v_2, \ldots, v_m$ be the indicator vectors of these clubs (defined in the same way as in the solution to Exercise 10.2.1). Thus, for each $i \in \{1, 2, \ldots, m\}$, the vector v_i is a row vector in $\mathbb{R}^n$ whose j-th coordinate is

$$\begin{cases} 1, & \text{if } j \in C_i; \\ 0, & \text{otherwise} \end{cases} \qquad \text{for each } j \in \{1, 2, \ldots, n\}.$$

These indicator vectors $v_1, v_2, \ldots, v_m$ are distinct (since no two clubs have the exact same set of members).

Recall one more concept from linear algebra: The *dot product* of two vectors $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and $y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$ is defined to be the scalar

$$x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{p=1}^{n} x_p y_p.$$

This dot product is denoted by $\langle x, y \rangle$ or sometimes by $x \cdot y$. We will use the notation $\langle x, y \rangle$ in the following.

For any $i, j \in \{1, 2, \ldots, m\}$, we have

$$\left\langle v_i, v_j \right\rangle = \sum_{p=1}^{n} \underbrace{\begin{cases} 1, & \text{if } p \in C_i; \\ 0, & \text{otherwise} \end{cases} \cdot \begin{cases} 1, & \text{if } p \in C_j; \\ 0, & \text{otherwise} \end{cases}}_{= \begin{cases} 1, & \text{if } p \in C_i \text{ and } p \in C_j; \\ 0, & \text{otherwise} \end{cases}}$$

$$\begin{pmatrix} \text{since } 1 \cdot 1 = 1 \text{ but } 1 \cdot 0 = 0 \cdot 1 = 0 \cdot 0 = 0 \end{pmatrix}$$

$$\begin{pmatrix} \text{by the definition of the dot product} \\ \text{and of the indicator vectors } v_i \text{ and } v_j \end{pmatrix}$$

$$= \sum_{p=1}^{n} \begin{cases} 1, & \text{if } p \in C_i \text{ and } p \in C_j; \\ 0, & \text{otherwise} \end{cases} = \sum_{p=1}^{n} \begin{cases} 1, & \text{if } p \in C_i \cap C_j; \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{pmatrix} \text{since the statement "} p \in C_i \text{ and } p \in C_j \text{"} \\ \text{is equivalent to "} p \in C_i \cap C_j \text{"} \end{pmatrix}$$

$$= \left| C_i \cap C_j \right| \tag{4}$$

(since this sum has a nonzero addend corresponding to each $p \in C_i \cap C_j$, and all these nonzero addends equal 1).

Therefore, for any two **distinct** elements $i, j \in \{1, 2, \ldots, m\}$, we have

$$\left\langle v_i, v_j \right\rangle = \left| C_i \cap C_j \right| = k \tag{5}$$

(by our assumption that any two distinct clubs share exactly k members). On the other hand, for each $i \in \{1, 2, \ldots, m\}$, we have

$$\begin{aligned} \left\langle v_i, v_i \right\rangle &= \left| C_i \cap C_i \right| \qquad \text{(by (4), applied to } j = i\text{)} \\ &= \left| C_i \right|. \end{aligned} \tag{6}$$

Now, we must prove that $m \leq n$. Assume the contrary. Thus, $m > n$. Hence, by Theorem 10.2.1 **(a)**, the m vectors $v_1, v_2, \ldots, v_m$ in the n-dimensional vector space $\mathbb{R}^n$ are linearly dependent. In other words, there exist reals $c_1, c_2, \ldots, c_m$, not all zero, such that

$$c_1 v_1 + c_2 v_2 + \cdots + c_m v_m = \mathbf{0}$$

(where $\mathbf{0}$ denotes the zero vector). Consider these reals. Thus,

$$\mathbf{0} = c_1 v_1 + c_2 v_2 + \cdots + c_m v_m = \sum_i c_i v_i.$$

(Here and in the following, all summation indices range from 1 to m. Thus, "$\sum_i$" means "$\sum_{i=1}^{m}$", and similarly for any other sums.)

It is easy to see that each $i \in \{1, 2, \ldots, m\}$ satisfies

$$\left| C_i \right| \geq k. \tag{7}$$

(Indeed, if $i \in \{1,2,\ldots,m\}$, then we can pick any arbitrary $j \in \{1,2,\ldots,m\}$ that is distinct from i [8], and then we have $C_i \supseteq C_i \cap C_j$ and therefore $|C_i| \geq |C_i \cap C_j| = k$ (by our assumption that any two distinct clubs share exactly k members). Thus, (7) is proved.)

Now, from $\mathbf{0} = \sum_i c_i v_i$ and $\mathbf{0} = \sum_i c_i v_i = \sum_j c_j v_j$, we obtain

$$\begin{aligned}
\langle \mathbf{0}, \mathbf{0} \rangle &= \left\langle \sum_i c_i v_i, \ \sum_j c_j v_j \right\rangle \\
&= \sum_i \sum_j c_i c_j \left\langle v_i, v_j \right\rangle \qquad \left(\begin{array}{c} \text{since the dot product is linear} \\ \text{in each of its two arguments} \end{array} \right) \\
&= \sum_{i,j} c_i c_j \left\langle v_i, v_j \right\rangle \qquad \left(\begin{array}{c} \text{of course, both indices } i \text{ and } j \\ \text{range over } \{1,2,\ldots,m\} \end{array} \right) \\
&= \underbrace{\sum_{\substack{i,j; \\ i=j}} c_i c_j \left\langle v_i, v_j \right\rangle}_{= \sum_i c_i c_i \langle v_i, v_i \rangle} + \sum_{\substack{i,j; \\ i \neq j}} c_i c_j \underbrace{\left\langle v_i, v_j \right\rangle}_{\substack{=k \\ \text{(by (5))}}} \\
&= \sum_i \underbrace{c_i c_i}_{=c_i^2} \underbrace{\left\langle v_i, v_i \right\rangle}_{\substack{=|C_i| \\ \text{(by (6))}}} + \underbrace{\sum_{\substack{i,j; \\ i \neq j}} c_i c_j k}_{= k \sum\limits_{\substack{i,j; \\ i \neq j}} c_i c_j} \\
&= \sum_i c_i^2 \left| C_i \right| + k \sum_{\substack{i,j; \\ i \neq j}} c_i c_j .
\end{aligned}$$

Comparing this with $\langle \mathbf{0}, \mathbf{0} \rangle = 0$ (which is obvious), we obtain

$$\sum_i c_i^2 \left| C_i \right| + k \sum_{\substack{i,j; \\ i \neq j}} c_i c_j = 0. \tag{8}$$

Now, let us simplify the second sum on the left hand side here. (It is generally a good idea to simplify the sums that have the most addends.) The trick is to observe

[8] Why does such a j exist? Because we have $m > n \geq 1$, and thus there are at least two different $j \in \{1,2,\ldots,m\}$ (and therefore at least one of them is distinct from i).

that

$$\left(\sum_i c_i\right)^2 = \left(\sum_i c_i\right)\left(\sum_i c_i\right) = \left(\sum_i c_i\right)\left(\sum_j c_j\right)$$
$$= \sum_{i,j} c_i c_j = \underbrace{\sum_{\substack{i,j;\\ i=j}} c_i c_j}_{=\sum_i c_i^2} + \sum_{\substack{i,j;\\ i\neq j}} c_i c_j = \sum_i c_i^2 + \sum_{\substack{i,j;\\ i\neq j}} c_i c_j,$$

so that

$$\sum_{\substack{i,j;\\ i\neq j}} c_i c_j = \left(\sum_i c_i\right)^2 - \sum_i c_i^2.$$

Thus,

$$\sum_i c_i^2 |C_i| + k \underbrace{\sum_{\substack{i,j;\\ i\neq j}} c_i c_j}_{=\left(\sum_i c_i\right)^2 - \sum_i c_i^2} = \sum_i c_i^2 |C_i| + k\left(\left(\sum_i c_i\right)^2 - \sum_i c_i^2\right)$$
$$= \sum_i c_i^2 \left(|C_i| - k\right) + k\left(\sum_i c_i\right)^2.$$

Comparing this with (8), we obtain

$$0 = \sum_i c_i^2 \left(|C_i| - k\right) + k\left(\sum_i c_i\right)^2. \tag{9}$$

This equality is peculiar in that the left hand side is 0, whereas the right hand side is a sum of nonnegative addends (indeed, all the addends $\underbrace{c_i^2}_{\substack{\geq 0\\ \text{(since squares are } \geq 0)}}$ $\underbrace{\left(|C_i| - k\right)}_{\substack{\geq 0\\ \text{(by (7))}}}$ and $\underbrace{k}_{>0}$ $\underbrace{\left(\sum_i c_i\right)^2}_{\substack{\geq 0\\ \text{(since squares are } \geq 0)}}$ on the right hand side are nonnegative). However, a sum of nonnegative reals can only be 0 if all addends are 0. Thus, the equality (9) entails that all the addends on the right hand side are 0. In other words,

$$c_i^2 \left(|C_i| - k\right) = 0 \qquad \text{for each } i \in \{1, 2, \ldots, m\}, \tag{10}$$

and

$$k\left(\sum_i c_i\right)^2 = 0. \tag{11}$$

From (11), we obtain

$$\sum_i c_i = 0 \tag{12}$$

(since k is nonzero). Since the reals $c_1, c_2, \ldots, c_m$ are not all zero, this entails that there exist **at least two** numbers $p \in \{1, 2, \ldots, m\}$ satisfying $c_p \neq 0$ (because if there was only one such number p, then the sum $\sum_i c_i$ would equal its single nonzero addend $c_p \neq 0$; but this would contradict (12)). In other words, there exist two distinct numbers i and j in $\{1, 2, \ldots, m\}$ that satisfy $c_i \neq 0$ and $c_j \neq 0$. Consider these i and j. From (10), we obtain $c_i^2 \left(|C_i| - k\right) = 0$. Since $c_i \neq 0$, we can divide this equality by c_i^2, and thus obtain $|C_i| - k = 0$. In other words, $|C_i| = k$. However, $\left|C_i \cap C_j\right| = k$ (by our assumption that any two distinct clubs share exactly k members). Now, $C_i \cap C_j$ is a subset of the finite set C_i that has the same size as C_i (because $\left|C_i \cap C_j\right| = k = |C_i|$). But the only such subset is obviously C_i. Thus, $C_i \cap C_j = C_i$. Similarly, $C_i \cap C_j = C_j$. Comparing these two equalities, we obtain $C_i = C_j$. But this contradicts the assumption that no two clubs have the exact same set of members. Hence, we found a contradiction, and the exercise is solved. □

Theorem 10.2.1 also leads to a useful general result about dimensions of vector spaces (similar to the well-known fact in combinatorics that any subset of an n-element set has size $\leq n$):

Theorem 10.2.7. Let W be a subspace of a finite-dimensional vector space V. (Here and in the following, "subspace" means "vector subspace".) Then:

(a) We have $\dim W \leq \dim V$.

(b) If $\dim W = \dim V$, then $V = W$.

Proof sketch. Set $n := \dim V$. Let $\mathbb{F}$ be the field over which V is a vector space.

The vector space V is n-dimensional (since $n = \dim V$). Thus, Theorem 10.2.1 **(a)** shows that V does not contain more than n linearly independent vectors. Hence, W does not contain more than n linearly independent vectors either (since W is a subspace of V). But of course, W contains 0 linearly independent vectors. Thus, there exists a **largest** integer $m \in \mathbb{N}$ such that W contains m linearly independent vectors. Consider this largest m, and pick a list $(w_1, w_2, \ldots, w_m)$ of m linearly independent vectors in W. Note that m cannot be larger than n (since W contains m linearly independent vectors but does not contain more than n linearly independent vectors). In other words, $m \leq n$. Also note that W does not contain $m + 1$

linearly independent vectors[9].

Now, we claim that the list $(w_1, w_2, \ldots, w_m)$ is a basis of W. Indeed, we already know that this list is linearly independent, so we only need to prove that it spans W. To this purpose, we fix a vector $w \in W$. Then, the list $(w_1, w_2, \ldots, w_m, w)$ consists of $m+1$ vectors, and thus cannot be linearly independent (because if it was, then W would contain $m+1$ linearly independent vectors, but we know that W does not contain $m+1$ linearly independent vectors). Hence, there exist scalars $c_1, c_2, \ldots, c_m, c \in \mathbb{F}$, not all zero, such that

$$c_1 w_1 + c_2 w_2 + \cdots + c_m w_m + cw = \mathbf{0} \tag{13}$$

(where $\mathbf{0}$ denotes the zero vector of V). Consider these scalars. If c was 0, then the equality (13) would simplify to $c_1 w_1 + c_2 w_2 + \cdots + c_m w_m = \mathbf{0}$ (since $\underbrace{c}_{=0} w = 0w = \mathbf{0}$), which would entail $c_1 = c_2 = \cdots = c_m = 0$ (since the list $(w_1, w_2, \ldots, w_m)$ is linearly independent) and therefore $c_1 = c_2 = \cdots = c_m = c = 0$ (since $c = 0$); but this would contradict the fact that $c_1, c_2, \ldots, c_m, c$ are not all zero. Hence, c cannot be 0. Thus, we can divide by c in the field $\mathbb{F}$. Dividing both sides of the equality (13) by c (that is, multiplying them by $\dfrac{1}{c}$), we obtain

$$\frac{c_1}{c} w_1 + \frac{c_2}{c} w_2 + \cdots + \frac{c_m}{c} w_m + w = \mathbf{0}.$$

Solving this for w, we find

$$\begin{aligned} w &= -\left(\frac{c_1}{c} w_1 + \frac{c_2}{c} w_2 + \cdots + \frac{c_m}{c} w_m \right) \\ &= \frac{-c_1}{c} w_1 + \frac{-c_2}{c} w_2 + \cdots + \frac{-c_m}{c} w_m \\ &\in \operatorname{span}\{w_1, w_2, \ldots, w_m\}. \end{aligned}$$

Forget that we fixed w. We thus have shown that $w \in \operatorname{span}\{w_1, w_2, \ldots, w_m\}$ for each $w \in W$. In other words, $W \subseteq \operatorname{span}\{w_1, w_2, \ldots, w_m\}$. Thus, the list $(w_1, w_2, \ldots, w_m)$ spans the vector space W. Since this list is also linearly independent, we conclude that it is a basis of W.

Hence, the vector space W has a basis consisting of m vectors (namely, this list $(w_1, w_2, \ldots, w_m)$). Thus, its dimension is $\dim W = m$. Now, $\dim W = m \leq n = \dim V$, so that Theorem 10.2.7 **(a)** is proved.

(b) Assume that $\dim W = \dim V$. Then, $m = \dim W = \dim V = n$. Now, recall that the m vectors $w_1, w_2, \ldots, w_m$ are linearly independent. In other words, the n vectors $w_1, w_2, \ldots, w_n$ are linearly independent (since $m = n$). Hence, $w_1, w_2, \ldots, w_n$ are n linearly independent vectors in the n-dimensional vector space V, and thus

[9] since m was chosen to be the **largest** integer with the property that W contains m linearly independent vectors

form a basis of V (by Theorem 10.2.1 **(b)**). Thus,

$$\begin{aligned} V &= \operatorname{span}\{w_1, w_2, \ldots, w_n\} = \operatorname{span}\{w_1, w_2, \ldots, w_m\} \qquad (\text{since } n = m) \\ &= W \qquad (\text{since the list } (w_1, w_2, \ldots, w_m) \text{ spans } W). \end{aligned}$$

This proves Theorem 10.2.7 **(b)**. □

10.3. Determinants

A highly useful subfield of linear algebra is the evaluation of determinants. This is an art all in itself, with a long history predating most of linear algebra (determinants were introduced two centuries before matrices![10]). An introduction to this art can be found in [Grinbe21, §6.4], more advanced methods in [Kratte99], and a multitude of exercises in [Prasol94, Chapter 1] and [Grinbe15, Chapter 6]. We will only give a few examples and applications here.

10.3.1. Definition and methods of computation

We recall that the determinant of an $n \times n$-matrix

$$A = \left(a_{i,j}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

is defined by the formula

$$\det A = \sum_{\sigma \in S_n} (-1)^{\sigma} a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}, \tag{14}$$

where

- the set S_n is the set of all permutations of the set $\{1, 2, \ldots, n\}$ (that is, bijective maps from this set $\{1, 2, \ldots, n\}$ to itself);
- the factor $(-1)^{\sigma}$ denotes the sign of the permutation σ (that is, 1 if σ is even, and -1 if σ is odd). (See, e.g., [Strick21, Appendix B] for a quick but well-written introduction to signs of permutations.)

[10] The word "matrix" was introduced by J. J. Sylvester to refer to a rectangular table of numbers out of which many determinants ("minor") can be formed by removing some rows and some columns ("as from the womb of a common parent"; thus the word "matrix"). Before Sylvester, authors would just talk about determinants $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ without having a special name for the table of entries "within" the determinant.

The above formula (14) is known as the *Leibniz formula*, and is probably the simplest way to define determinants, but one of the least efficient ways to compute them. Over the centuries, many techniques for the computation of determinants have appeared, such as the following:

- *Laplace expansion* (along a row or a column). See [Grinbe21, §6.4.7], [Grinbe15, §6.12] or [Strick21, Propositions B.24 and B.25] (or various other sources) for statements and proofs. For example, here is how a 3×3-determinant can be expanded along the second row and the second column:

  $$\det \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = -a' \det \begin{pmatrix} b & c \\ b'' & c'' \end{pmatrix} + b' \det \begin{pmatrix} a & c \\ a'' & c'' \end{pmatrix} - c' \det \begin{pmatrix} a & b \\ a'' & b'' \end{pmatrix}$$

  and

  $$\det \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = -b \det \begin{pmatrix} a' & c' \\ a'' & c'' \end{pmatrix} + b' \det \begin{pmatrix} a & c \\ a'' & c'' \end{pmatrix} - b'' \det \begin{pmatrix} a & c \\ a' & c' \end{pmatrix}.$$

  Note the signs, which always alternate and are positive whenever the entry being plucked out comes from the diagonal of the matrix! Tactically, Laplace expansion is at its most useful when many addends in the resulting sum vanish. For example, when computing $\det \begin{pmatrix} 1 & 0 & 2 \\ 2 & 3 & 0 \\ 4 & 0 & 0 \end{pmatrix}$ by Laplace expansion, it is reasonable to expand along the first row (since it has a 0 entry), but it is best to expand along the third row or the second or third column (since it has two 0 entries).

- *Multilinearity* and *alternatingness*. These basic properties of determinants (see [Grinbe21, Theorems 6.4.12 and 6.4.14] or [Grinbe15, Exercise 6.7], for example) look harmless, but are often helpful as steps in longer computations. Here is a quick summary:
  - If a row of a matrix A is zero (i.e., consists entirely of zeroes), then $\det A = 0$.
  - If two rows of a matrix A are equal, then $\det A = 0$.
  - If we swap two rows of a matrix A, then $\det A$ gets multiplied by -1.
  - If we multiply a row of a matrix A by a given scalar λ, then $\det A$ gets multiplied by λ.
  - If we add a scalar multiple of a row of a matrix A to another row of A, then $\det A$ remains unchanged.

– Let $k \in \{1, 2, \ldots, n\}$ be arbitrary. If three $n \times n$-matrices A, B and C satisfy

$$(\text{the } k\text{-th row of } C) = (\text{the } k\text{-th row of } A) + (\text{the } k\text{-th row of } B)$$

and

$$\begin{aligned}(\text{the } i\text{-th row of } C) &= (\text{the } i\text{-th row of } A) \\ &= (\text{the } i\text{-th row of } B) \qquad \text{for all } i \neq k,\end{aligned}$$

then

$$\det C = \det A + \det B.$$

This property (known as *multilinearity*) is best illustrated on an example (with $n = 3$ and $k = 2$):

$$\det \underbrace{\begin{pmatrix} a & b & c \\ d+d' & e+e' & f+f' \\ g & h & i \end{pmatrix}}_{\text{this is } C} = \det \underbrace{\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}}_{\text{this is } A} + \det \underbrace{\begin{pmatrix} a & b & c \\ d' & e' & f' \\ g & h & i \end{pmatrix}}_{\text{this is } B}.$$

– All of the above holds if "row" is replaced by "column".

- *Gaussian elimination*, and (more generally) artful use of *row and column operations*. This relies on the properties listed above, and is particular convenient when you have a fully specified matrix (with given size and given entries) in front of you, but sometimes works in more general settings.
- The transpose A^T of a matrix A has the same determinant as A (that is, $\det\left(A^T\right) = \det A$).
- The determinant of a triangular matrix equals the product of its diagonal entries. This holds both for upper-triangular and lower-triangular matrices. For example, for 3×3-matrices, this is saying that

$$\det \begin{pmatrix} a & b & c \\ 0 & b' & c' \\ 0 & 0 & c'' \end{pmatrix} = \det \begin{pmatrix} a & 0 & 0 \\ a' & b' & 0 \\ a'' & b'' & c'' \end{pmatrix} = ab'c''.$$

- *Multiplicativity*: i.e., the formula $\det(AB) = \det A \cdot \det B$ that holds for any two $n \times n$-matrices A and B. This is useful whenever you can factor your matrix as a product of two (or more) simpler matrices. Particularly helpful are factorizations into triangular matrices (such as LU-factorization), whose determinants can be immediately computed according to the previous bullet point. See [Grinbe21, §6.4.5] or [Kratte99, §2.6] for examples of this strategy in action.

Keep in mind that there is no simple formula for $\det(A+B)$ (it rarely ever equals $\det A + \det B$), and don't forget that $\det(\lambda A)$ for a number λ is not $\lambda \det A$ but $\lambda^n \det A$.

Note that there is also a formula for $\det(AB)$ when A and B are not square (but AB is). This is the *Cauchy–Binet formula* ([Grinbe21, §6.4.3]), and is less simple than the $n \times n$-case (of course, $\det A$ and $\det B$ are not defined when A and B are not square, so it cannot possibly be as simple).

- *Eigenvalues*. If the n eigenvalues of an $n \times n$-matrix A are $\lambda_1, \lambda_2, \ldots, \lambda_n$ (listed with their algebraic multiplicities), then $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$. This fact is not very helpful for computing $\det A$, since any practical use requires computing all eigenvalues of A without going through the characteristic polynomial in the first place (since the latter is no easier than computing $\det A$). Nevertheless, there have been situations in which this has been useful (e.g., [Zhao09, Problem 3] or [Grinbe23b, §5.14.5 and §5.15.2]).

- The *invertible matrix theorem* says (among many other things) that an $n \times n$-matrix A has determinant 0 if and only if it has a nonzero nullspace (i.e., if there exists a nonzero vector v such that $Av = \mathbf{0}$). This can be used to prove that some determinants are 0.

- *Factor hunting* (aka *identification of factors*). This works when the entries of your matrix are polynomials, ideally in several indeterminates. See [Grinbe21, §6.4.6] or [Kratte99, §2.4].

- *Block matrices* and *Schur complements*. The most useful facts here are (I am assuming that you are familiar with block matrix notation)
  - the formula $\det \begin{pmatrix} A & B \\ \mathbf{0} & D \end{pmatrix} = \det A \cdot \det D$, where the matrix on the left is a block matrix with square blocks A and D. (The $\mathbf{0}$ in the bottom left means a zero matrix.)
  - the Schur complement formula $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \cdot \det\left(D - CA^{-1}B\right)$, where the matrix on the left is a block matrix with square blocks A and D. This requires A to be invertible.

- The *Sylvester identity* (one of the many): If A is an $n \times m$-matrix, and if B is an $m \times n$-matrix, then $\det(I_n + AB) = \det(I_m + BA)$. See [Zhao09, Remark after Problem 5] for a sketch of another proof.

  The usefulness of this identity comes mainly from the fact that n can be much larger than m, in which case it reduces a large determinant (in terms of matrix size) to a small one. A striking example of this trick in use is found in the second solution to problem B5 in the 60th Putnam contest 1999 on Kiran Kedlaya's website.

- *Dodgson condensation*. This is a recursive method based upon a surprising formula. See [Grinbe21, §6.4.8] and [Kratte99, §2.3] for details.

10.3.2. Computing practice

Here is an example of a matrix whose determinant can be computed using some of the above techniques.

Exercise 10.3.1. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n, b$ be $n+1$ numbers (e.g., real or complex). Compute the determinant

$$\det \underbrace{\begin{pmatrix} a_1+b & b & b & \cdots & b \\ b & a_2+b & b & \cdots & b \\ b & b & a_3+b & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a_n+b \end{pmatrix}}_{\substack{\text{This is the } n \times n\text{-matrix whose} \\ \text{diagonal entries are } a_1+b,\ a_2+b,\ \ldots,\ a_n+b \\ \text{and whose off-diagonal entries all equal } b}}.$$

First solution. The cases $n = 0$ and $n = 1$ are easily solved by hand: The determinant is 1 if $n = 0$ (since the 0×0-matrix has determinant 1), and is $a_1 + b$ if $n = 1$.

Now, assume that $n \geq 2$. Let us denote our matrix (i.e., the $n \times n$-matrix whose diagonal entries are $a_1 + b,\ a_2 + b,\ \ldots,\ a_n + b$ and whose off-diagonal entries all equal b) by $A(a_1, a_2, \ldots, a_n)$ (in order to stress its dependence on $a_1, a_2, \ldots, a_n$ [11]). Thus, we must compute $\det(A(a_1, a_2, \ldots, a_n))$.

We recall that the determinant of a matrix does not change when we add a scalar multiple of a row to another row. In particular, it does not change if we subtract a row from another row. Therefore, the determinant of the matrix

$$A(a_1, a_2, \ldots, a_n) = \begin{pmatrix} a_1+b & b & b & \cdots & b & b \\ b & a_2+b & b & \cdots & b & b \\ b & b & a_3+b & \cdots & b & b \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & b & \cdots & a_{n-1}+b & b \\ b & b & b & \cdots & b & a_n+b \end{pmatrix}$$

does not change when we subtract the second-to-last row from the last row. We

[11] Of course, it also depends on b, but we have no need to make this explicit in our notation, since b will remain unchanged in our entire solution.

thus obtain

$$\begin{aligned}
&\det\left(A\left(a_1, a_2, \ldots, a_n\right)\right) \\
&= \det\begin{pmatrix} a_1+b & b & b & \cdots & b & b \\ b & a_2+b & b & \cdots & b & b \\ b & b & a_3+b & \cdots & b & b \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & b & \cdots & a_{n-1}+b & b \\ b-b & b-b & b-b & \cdots & b-\left(a_{n-1}+b\right) & \left(a_n+b\right)-b \end{pmatrix} \\
&= \det\begin{pmatrix} a_1+b & b & b & \cdots & b & b \\ b & a_2+b & b & \cdots & b & b \\ b & b & a_3+b & \cdots & b & b \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & b & \cdots & a_{n-1}+b & b \\ 0 & 0 & 0 & \cdots & -a_{n-1} & a_n \end{pmatrix}.
\end{aligned}$$

The matrix on the right hand side has the nice property that the first $n-2$ entries of its last row are 0's. This suggests computing its determinant by Laplace expansion along the last row. In this expansion, all but the last two addends vanish (since the

0's from the last row appear as factors), and we are left with

$$\det\begin{pmatrix} a_1+b & b & b & \cdots & b & b \\ b & a_2+b & b & \cdots & b & b \\ b & b & a_3+b & \cdots & b & b \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & b & \cdots & a_{n-1}+b & b \\ 0 & 0 & 0 & \cdots & -a_{n-1} & a_n \end{pmatrix}$$

$$= \underbrace{(-1)^{n+(n-1)}}_{=-1}(-a_{n-1})\det\underbrace{\begin{pmatrix} a_1+b & b & b & \cdots & b & b \\ b & a_2+b & b & \cdots & b & b \\ b & b & a_3+b & \cdots & b & b \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & b & \cdots & a_{n-2}+b & b \\ b & b & b & \cdots & b & b \end{pmatrix}}_{=A(a_1,a_2,\ldots,a_{n-2},0)}$$

$$+ \underbrace{(-1)^{n+n}}_{=1} a_n \det\underbrace{\begin{pmatrix} a_1+b & b & b & \cdots & b \\ b & a_2+b & b & \cdots & b \\ b & b & a_3+b & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a_{n-1}+b \end{pmatrix}}_{=A(a_1,a_2,\ldots,a_{n-1})}$$

$$= \underbrace{(-1)(-a_{n-1})}_{=a_{n-1}}\det\left(A\left(a_1,a_2,\ldots,a_{n-2},0\right)\right) + a_n\det\left(A\left(a_1,a_2,\ldots,a_{n-1}\right)\right)$$

$$= a_{n-1}\det\left(A\left(a_1,a_2,\ldots,a_{n-2},0\right)\right) + a_n\det\left(A\left(a_1,a_2,\ldots,a_{n-1}\right)\right).$$

Altogether, we thus have

$$\det\left(A\left(a_1,a_2,\ldots,a_n\right)\right) = \det\begin{pmatrix} a_1+b & b & b & \cdots & b & b \\ b & a_2+b & b & \cdots & b & b \\ b & b & a_3+b & \cdots & b & b \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b & b & b & \cdots & a_{n-1}+b & b \\ 0 & 0 & 0 & \cdots & -a_{n-1} & a_n \end{pmatrix}$$

$$= a_{n-1}\det\left(A\left(a_1,a_2,\ldots,a_{n-2},0\right)\right) + a_n\det\left(A\left(a_1,a_2,\ldots,a_{n-1}\right)\right). \tag{15}$$

This is a recursive formula for $\det\left(A\left(a_1,a_2,\ldots,a_n\right)\right)$ that allows us to easily compute these determinants for small values of n, starting with

$$\det\left(A\left(\right)\right) = 1 \qquad \text{(this was the case } n=0\text{)} \qquad \text{and}$$
$$\det\left(A\left(a_1\right)\right) = a_1+b \qquad \text{(this was the case } n=1\text{)}.$$

We obtain

$$\begin{aligned}\det\left(A\left(a_1,a_2\right)\right) &= a_1 \underbrace{\det\left(A\left(0\right)\right)}_{=0+b} + a_2 \underbrace{\det\left(A\left(a_1\right)\right)}_{=a_1+b} \\ &= a_1\left(0+b\right) + a_2\left(a_1+b\right) \\ &= a_1a_2 + b\left(a_1+a_2\right)\end{aligned}$$

and similarly

$$\det\left(A\left(a_1,a_2,a_3\right)\right) = a_1a_2a_3 + b\left(a_1a_2+a_1a_3+a_2a_3\right)$$

and

$$\det\left(A\left(a_1,a_2,a_3,a_4\right)\right) = a_1a_2a_3a_4 + b\left(a_1a_2a_3+a_1a_2a_4+a_1a_3a_4+a_2a_3a_4\right).$$

This suggests a general formula: We claim that for every $n \in \mathbb{N}$ and every n numbers $a_1, a_2, \ldots, a_n$, we have

$$\begin{aligned}&\det\left(A\left(a_1,a_2,\ldots,a_n\right)\right) \\ &= a_1a_2\cdots a_n + b\sum_{i=1}^{n} a_1a_2\cdots\widehat{a_i}\cdots a_n,\end{aligned} \tag{16}$$

where the hat over the "a_i" means that the factor a_i is being omitted from the product (i.e., the product $a_1a_2\cdots\widehat{a_i}\cdots a_n$ should be read as $a_1a_2\cdots a_{i-1}a_{i+1}a_{i+2}\cdots a_n$).

It remains to prove our formula (16). This can be done quite easily by induction on n: The cases $n=0$ and $n=1$ are obvious and can be used as *base cases*. For the *induction step*, we fix an integer $n \geq 2$, and we assume (as the induction hypothesis) that (16) has already been proved for $n-1$ instead of n. Now, we must prove that (16) also holds for n. Since $n \geq 2$, we can use the formula (15). Our induction hypothesis yields that (16) holds for $n-1$ instead of n. In other words, the equality

$$\det\left(A\left(a_1,a_2,\ldots,a_{n-1}\right)\right) = a_1a_2\cdots a_{n-1} + b\sum_{i=1}^{n-1} a_1a_2\cdots\widehat{a_i}\cdots a_{n-1}$$

holds. Applying this same equality to 0 instead of a_{n-1}, we obtain

$$\det\left(A\left(a_1,a_2,\ldots,a_{n-2},0\right)\right) = a_1a_2\cdots a_{n-2}0 + b\sum_{i=1}^{n-1} a_1a_2\cdots\widehat{a_i}\cdots a_{n-2}0,$$

where the expression "$a_1a_2\cdots\widehat{a_i}\cdots a_{n-2}0$" is to be read as the product $a_1a_2\cdots a_{n-2}0$ with its i-th factor removed. In view of

$$\begin{aligned}\sum_{i=1}^{n-1} a_1a_2\cdots\widehat{a_i}\cdots a_{n-2}0 &= \sum_{i=1}^{n-2} \underbrace{a_1a_2\cdots\widehat{a_i}\cdots a_{n-2}0}_{\substack{=0 \\ \text{(since the last factor of this product is 0)}}} + \underbrace{a_1a_2\cdots a_{n-2}\widehat{0}}_{=a_1a_2\cdots a_{n-2}} \\ &\qquad \left(\begin{array}{c}\text{here, we have removed the addend} \\ \text{for } i=n-1 \text{ from the sum}\end{array}\right) \\ &= \sum_{i=1}^{n-2} 0 + a_1a_2\cdots a_{n-2} = a_1a_2\cdots a_{n-2},\end{aligned}$$

this becomes

$$\det\left(A\left(a_1, a_2, \ldots, a_{n-2}, 0\right)\right) = \underbrace{a_1 a_2 \cdots a_{n-2} 0}_{=0} + b \underbrace{\sum_{i=1}^{n-1} a_1 a_2 \cdots \widehat{a_i} \cdots a_{n-2} 0}_{=a_1 a_2 \cdots a_{n-2}}$$
$$= b \cdot a_1 a_2 \cdots a_{n-2}.$$

Plugging these results into (15), we find

$$\begin{aligned}
&\det\left(A\left(a_1, a_2, \ldots, a_n\right)\right) \\
&= a_{n-1} \underbrace{\det\left(A\left(a_1, a_2, \ldots, a_{n-2}, 0\right)\right)}_{=b \cdot a_1 a_2 \cdots a_{n-2}} + a_n \underbrace{\det\left(A\left(a_1, a_2, \ldots, a_{n-1}\right)\right)}_{=a_1 a_2 \cdots a_{n-1} + b \sum\limits_{i=1}^{n-1} a_1 a_2 \cdots \widehat{a_i} \cdots a_{n-1}} \\
&= \underbrace{a_{n-1} \cdot b \cdot a_1 a_2 \cdots a_{n-2}}_{\substack{=b \cdot a_1 a_2 \cdots a_{n-2} \cdot a_{n-1} \\ =b \cdot a_1 a_2 \cdots a_{n-1}}} + \underbrace{a_n \cdot \left(a_1 a_2 \cdots a_{n-1} + b \sum_{i=1}^{n-1} a_1 a_2 \cdots \widehat{a_i} \cdots a_{n-1}\right)}_{=a_n \cdot a_1 a_2 \cdots a_{n-1} + a_n \cdot b \sum\limits_{i=1}^{n-1} a_1 a_2 \cdots \widehat{a_i} \cdots a_{n-1}} \\
&= b \cdot a_1 a_2 \cdots a_{n-1} + \underbrace{a_n \cdot a_1 a_2 \cdots a_{n-1}}_{\substack{=a_1 a_2 \cdots a_{n-1} \cdot a_n \\ =a_1 a_2 \cdots a_n}} + \underbrace{a_n \cdot b \sum_{i=1}^{n-1} a_1 a_2 \cdots \widehat{a_i} \cdots a_{n-1}}_{=b \sum\limits_{i=1}^{n-1} a_1 a_2 \cdots \widehat{a_i} \cdots a_{n-1} \cdot a_n} \\
&= b \cdot a_1 a_2 \cdots a_{n-1} + a_1 a_2 \cdots a_n + b \sum_{i=1}^{n-1} a_1 a_2 \cdots \widehat{a_i} \cdots a_{n-1} \cdot a_n.
\end{aligned}$$

Comparing this with

$$a_1 a_2 \cdots a_n + b \underbrace{\sum_{i=1}^{n} a_1 a_2 \cdots \widehat{a_i} \cdots a_n}_{=\sum\limits_{i=1}^{n-1} a_1 a_2 \cdots \widehat{a_i} \cdots a_n + a_1 a_2 \cdots \widehat{a_n}}$$

(here, we have split off the addend for $i=n$ from the sum)

$$\begin{aligned}
&= a_1 a_2 \cdots a_n + b \left(\sum_{i=1}^{n-1} a_1 a_2 \cdots \widehat{a_i} \cdots a_n + a_1 a_2 \cdots \widehat{a_n}\right) \\
&= a_1 a_2 \cdots a_n + b \sum_{i=1}^{n-1} \underbrace{a_1 a_2 \cdots \widehat{a_i} \cdots a_n}_{=a_1 a_2 \cdots \widehat{a_i} \cdots a_{n-1} \cdot a_n} + b \cdot \underbrace{a_1 a_2 \cdots \widehat{a_n}}_{=a_1 a_2 \cdots a_{n-1}} \\
&= a_1 a_2 \cdots a_n + b \sum_{i=1}^{n-1} a_1 a_2 \cdots \widehat{a_i} \cdots a_{n-1} \cdot a_n + b \cdot a_1 a_2 \cdots a_{n-1} \\
&= b \cdot a_1 a_2 \cdots a_{n-1} + a_1 a_2 \cdots a_n + b \sum_{i=1}^{n-1} a_1 a_2 \cdots \widehat{a_i} \cdots a_{n-1} \cdot a_n,
\end{aligned}$$

we obtain

$$\det\left(A\left(a_1, a_2, \ldots, a_n\right)\right) = a_1 a_2 \cdots a_n + b \sum_{i=1}^{n} a_1 a_2 \cdots \widehat{a_i} \cdots a_n.$$

In other words, (16) holds for n. This completes the induction step. Thus, (16) is proved, and the problem solved. □

Second solution (sketched). We start in a similar way as in our first solution, namely by recalling that the determinant of a matrix does not change when we add a scalar multiple of a row to another row. But this time, let us perform not one, but several such row operations: We subtract the first row from each of the other rows of our matrix. Thus, we obtain

$$\begin{aligned}
&\det\begin{pmatrix} a_1+b & b & b & \cdots & b \\ b & a_2+b & b & \cdots & b \\ b & b & a_3+b & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a_n+b \end{pmatrix} \\
&= \det\begin{pmatrix} a_1+b & b & b & \cdots & b \\ -a_1 & a_2 & 0 & \cdots & 0 \\ -a_1 & 0 & a_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_1 & 0 & 0 & \cdots & a_n \end{pmatrix}.
\end{aligned}$$

This has noticeably simplified our matrix, at least in the sense that it now has lots of zero entries – but we still don't have an immediate way to compute its determinant.[12] So let us perform some more row operations.

We can try to get rid of the b's in the first row by subtracting appropriate multiples of other rows. Namely, let us subtract the $\dfrac{b}{a_i}$-multiple of the i-th row from the first row for each $i \in \{2, 3, \ldots, n\}$. The factor $\dfrac{b}{a_i}$ was chosen here precisely to cancel the b in the first row (as in Gaussian elimination); we obtain the matrix

$$\begin{pmatrix} r & 0 & 0 & \cdots & 0 \\ -a_1 & a_2 & 0 & \cdots & 0 \\ -a_1 & 0 & a_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_1 & 0 & 0 & \cdots & a_n \end{pmatrix},$$

where

$$r = (a_1 + b) - \sum_{i=2}^{n} \frac{b}{a_i}\left(-a_1\right).$$

[12] Actually, we could compute it immediately if we knew the formula for determinants of arrowhead matrices ([Grinbe18, Exercise 6 **(b)**]). But let us try to avoid any apocryphal formulas here.

This matrix is lower-triangular, and thus its determinant is the product of its diagonal entries, i.e., the product

$$ra_2a_3\cdots a_n.$$

Since all our row operations left the determinant of the matrix unchanged, we thus conclude that the original matrix must have determinant $ra_2a_3\cdots a_n$ as well.

Are we done? Almost. We have tacitly assumed that the numbers $a_2, a_3, \ldots, a_n$ are nonzero, since we have been dividing by them. Even our final result $ra_2a_3\cdots a_n$ makes no sense if one of these numbers is zero, since these numbers appear as denominators in the definition of r. Thus, we have solved the problem in the "generic" case when the numbers $a_2, a_3, \ldots, a_n$ are nonzero, but the "exceptional" case when some of them are zero still remains to be addressed.

Let us address it in steps: First, we adapt the answer; then, we will adapt the proof. To adapt the answer, we rewrite our formula $ra_2a_3\cdots a_n$ by plugging the definition of r into it:

$$\begin{aligned}
ra_2a_3\cdots a_n &= \left((a_1+b) - \sum_{i=2}^{n} \frac{b}{a_i}(-a_1)\right) a_2a_3\cdots a_n \\
&= (a_1+b)\, a_2a_3\cdots a_n - \sum_{i=2}^{n} \underbrace{\frac{b}{a_i}(-a_1)\, a_2a_3\cdots a_n}_{\substack{= -\frac{b}{a_i}\cdot a_1a_2\cdots a_n \\ = -b\cdot\dfrac{a_1a_2\cdots a_n}{a_i} \\ = -ba_1a_2\cdots\widehat{a_i}\cdots a_n \\ \text{(where the hat over the "}a_i\text{" is} \\ \text{understood as in the First solution)}}} \\
&= (a_1+b)\, a_2a_3\cdots a_n - \sum_{i=2}^{n} \left(-ba_1a_2\cdots\widehat{a_i}\cdots a_n\right).
\end{aligned}$$

This latter expression involves no fractions any more, so it makes sense in both "generic" and "exceptional" cases. We can actually rewrite it in a nicer (more

symmetric) form:

$$\begin{aligned}
&(a_1+b)\,a_2a_3\cdots a_n - \sum_{i=2}^{n}\left(-ba_1a_2\cdots\widehat{a_i}\cdots a_n\right)\\
&= \underbrace{(a_1+b)\,a_2a_3\cdots a_n}_{=a_1\cdot a_2a_3\cdots a_n + b\cdot a_2a_3\cdots a_n} + \sum_{i=2}^{n} ba_1a_2\cdots\widehat{a_i}\cdots a_n\\
&= \underbrace{a_1\cdot a_2a_3\cdots a_n}_{=a_1a_2\cdots a_n} + \underbrace{b\cdot a_2a_3\cdots a_n + \sum_{i=2}^{n} ba_1a_2\cdots\widehat{a_i}\cdots a_n}_{=b\left(a_2a_3\cdots a_n + \sum\limits_{i=2}^{n} a_1a_2\cdots\widehat{a_i}\cdots a_n\right)}\\
&= a_1a_2\cdots a_n + b\underbrace{\left(a_2a_3\cdots a_n + \sum_{i=2}^{n} a_1a_2\cdots\widehat{a_i}\cdots a_n\right)}_{=\sum\limits_{i=1}^{n} a_1a_2\cdots\widehat{a_i}\cdots a_n}\\
&= a_1a_2\cdots a_n + b\sum_{i=1}^{n} a_1a_2\cdots\widehat{a_i}\cdots a_n.
\end{aligned}$$

Altogether, our claimed formula is therefore

$$\begin{aligned}
&\det\begin{pmatrix}
a_1+b & b & b & \cdots & b\\
b & a_2+b & b & \cdots & b\\
b & b & a_3+b & \cdots & b\\
\vdots & \vdots & \vdots & \ddots & \vdots\\
b & b & b & \cdots & a_n+b
\end{pmatrix}\\
&= a_1a_2\cdots a_n + b\sum_{i=1}^{n} a_1a_2\cdots\widehat{a_i}\cdots a_n. \qquad (17)
\end{aligned}$$

This is the exact same formula (16) that we obtained in the First solution. But (in order to get an independent second solution) we still need to prove it in the "exceptional" case. There are several ways to do this:

1. One way is to observe that (17) is a polynomial identity in $a_1, a_2, \ldots, a_n, b$ (that is, both of its sides are polynomial functions in each of the arguments $a_1, a_2, \ldots, a_n, b$). Thus, the polynomial identity trick (Corollary 8.5.7 on Worksheet 8) shows that (for instance) if it holds for infinitely many values of a_2, then it holds for all values of a_2 (where all the remaining arguments $a_1, a_3, a_4, \ldots, b$ are held constant). Thus, in particular, if it holds for all nonzero a_2, then it holds for all values of a_2 (including 0). By drawing such conclusion several times (once for each of the arguments $a_2, a_3, \ldots, a_n$), we see that if the equality (17) holds for all nonzero $a_2, a_3, \ldots, a_n$, then it holds for all $a_2, a_3, \ldots, a_n$ (including 0). In other words, if (17) holds in the "generic" case,

then it holds in all cases. Since we have already proved (17) in the "generic" case, this completes our proof.

2. Alternatively, we can prove (17) in the "exceptional" case by hand: First, we observe that $a_1, a_2, \ldots, a_n$ play symmetric roles in the equality (17) (since the determinant of a matrix responds in very simple ways when we permute its rows or columns). Hence, we can permute the numbers $a_1, a_2, \ldots, a_n$ at will[13]. In particular, if **exactly one** of these numbers $a_1, a_2, \ldots, a_n$ is 0, then we can permute them so that this number becomes a_1, while the remaining numbers $a_2, a_3, \ldots, a_n$ are nonzero; but this lands us in the "generic" case, and we already know that (17) holds in this case. Hence, it remains to only consider the case when **at least two** of the numbers $a_1, a_2, \ldots, a_n$ are 0. But this case is particularly simple: In this case, the left hand side of (17) is 0 because the matrix has two equal columns[14], whereas the right hand side of (17) is 0 because each of the products has at least one vanishing factor. All in all, (17) has thus been proved.

In either case, we are done. □

Yet another solution to Exercise 10.3.1 (with a slight change of notations: each a_i is replaced by $a_i - b$) can be found in `https://math.stackexchange.com/questions/2110766/a/2112473#2112473` .

More examples of matrices with nicely computable determinants can be found in the class and homework problems below.

10.3.3. A Putnam problem

Determinants have been originally introduced for solving systems of linear equations. Better ways are known for this nowadays, but determinants still have a theoretical significance in determining how many solutions a given system has.

Exercise 10.3.2. Let $n \in \mathbb{N}$. Assume that $a_1, a_2, \ldots, a_{2n+1}$ are $2n+1$ real numbers with the following property:

> *Splitting property:* If any of the $2n+1$ numbers $a_1, a_2, \ldots, a_{2n+1}$ is removed, then the remaining $2n$ numbers can be split into two equinumerous heaps with equal sum. ("Equinumerous" means that each heap contains exactly n numbers.)

Prove that all $2n+1$ numbers $a_1, a_2, \ldots, a_{2n+1}$ are equal.
(34th Putnam contest 1973, problem B1, generalized)

[13] In more details: Swapping two numbers a_i and a_j is tantamount to swapping the i-th row with the j-th row and then swapping the i-th column with the j-th column in the matrix on the left hand side of (17). The row swap multiplies the determinant by -1; then the column swap multiplies it by -1 again. Both operations combined therefore leave the determinant unchanged (since $(-1)(-1) = 1$). Thus, we can swap two of our numbers $a_1, a_2, \ldots, a_n$ without affecting the correctness of (17).

[14] Namely, if $a_i = 0$ and $a_j = 0$, then the i-th and j-th columns of the matrix are equal.

Solution idea. I shall proceed informally: I will show the argument on an example, while occasionally explaining why the argument generalizes.

For my example, I take $n = 2$. Thus, we have $2n+1 = 5$ real numbers a_1, a_2, a_3, a_4, a_5 satisfying the splitting property. The splitting property is saying that if we remove any of these 5 numbers, then the remaining 4 can be split into two equinumerous heaps (i.e., two heaps of 2 numbers each) with equal sum. For example, if a_1 is removed, then the remaining 4 numbers a_2, a_3, a_4, a_5 can be split into two such heaps – meaning that we have $a_2 + a_3 = a_4 + a_5$ or $a_2 + a_4 = a_3 + a_5$ or $a_3 + a_4 = a_2 + a_5$. For example, let me assume that $a_2 + a_4 = a_3 + a_5$. Likewise, I assume that removing a_2 leads to $a_1 + a_3 = a_4 + a_5$; that removing a_3 leads to $a_1 + a_5 = a_2 + a_4$; that removing a_4 leads to $a_1 + a_2 = a_3 + a_5$; and that removing a_5 leads to $a_1 + a_3 = a_2 + a_4$. Altogether, we now know that our five numbers a_1, a_2, a_3, a_4, a_5 satisfy the five linear equations

$$\begin{cases} a_2 + a_4 = a_3 + a_5; \\ a_1 + a_3 = a_4 + a_5; \\ a_1 + a_5 = a_2 + a_4; \\ a_1 + a_2 = a_3 + a_5; \\ a_1 + a_3 = a_2 + a_4. \end{cases} \tag{18}$$

Our goal is to prove that all 5 numbers a_1, a_2, a_3, a_4, a_5 are equal. We observe that if we subtract one and the same real number b from each of our 5 numbers a_1, a_2, a_3, a_4, a_5 (that is, if we replace each a_i by $a_i - b$), then nothing really changes: Our 5 numbers still satisfy the splitting property after this transformation (because the sum of the two numbers in either heap is decreased by $2b$), and our goal (to prove that a_1, a_2, a_3, a_4, a_5 are equal) remains the same.[15]

Thus, we can freely choose a real number b and subtract it from each of our 5 numbers a_1, a_2, a_3, a_4, a_5. Let us choose $b = a_5$. Then, this subtraction results in a_5 becoming $a_5 - a_5 = 0$. Hence, we can WLOG assume that $a_5 = 0$. Assume this. Of course, in the general case (as opposed to the example we are looking at), the assumption is $a_{2n+1} = 0$ rather than $a_5 = 0$, but it is obtained in the same way (viz., subtracting a_{2n+1} from each of our $2n+1$ numbers $a_1, a_2, \ldots, a_{2n+1}$).

Since $a_5 = 0$, we can simplify the system of equations (18) by removing each appearance of a_5. We thus obtain

$$\begin{cases} a_2 + a_4 = a_3; \\ a_1 + a_3 = a_4; \\ a_1 = a_2 + a_4; \\ a_1 + a_2 = a_3; \\ a_1 + a_3 = a_2 + a_4. \end{cases} \tag{19}$$

[15] Note that we are using the "equinumerous" requirement in the splitting property here! If the two heaps were not required to be equinumerous, then their sums could decrease by different multiples of b when we subtract b from each of our 5 numbers a_1, a_2, a_3, a_4, a_5.

Furthermore, let us remove the last equation from this system, thus obtaining

$$\begin{cases} a_2 + a_4 = a_3; \\ a_1 + a_3 = a_4; \\ a_1 = a_2 + a_4; \\ a_1 + a_2 = a_3. \end{cases} \tag{20}$$

Recall that our goal is to prove that all 5 numbers a_1, a_2, a_3, a_4, a_5 are equal. Equivalently, we must prove that all 4 numbers a_1, a_2, a_3, a_4 are 0 (since a_5 is already 0).

Thus, our goal is an instance of a well-known basic problem in linear algebra: To show that the only solution of a certain system of linear equations (specifically, (20)) is the zero vector (i.e., the solution where each unknown is 0).

The system (20) of linear equations can, of course, be solved by Gaussian elimination, but this would not be generalizable beyond the specific example I have chosen. So let us instead analyze it from a "bird's-eye view", without looking too closely at the specifics. As with any system of linear equations, we can rewrite it in the form "matrix times unknown vector equals known vector". Rewritten in this form, it becomes

$$\begin{pmatrix} 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & -1 \\ 1 & -1 & 0 & -1 \\ 1 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

In other words, it becomes

$$A \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{21}$$

where

$$A := \begin{pmatrix} 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & -1 \\ 1 & -1 & 0 & -1 \\ 1 & 1 & -1 & 0 \end{pmatrix}.$$

The entries of this matrix A are specific to our example, but the general structure is always the same:

- The matrix A has dimensions $2n \times 2n$.
- All diagonal entries of A are 0's, since the i-th equation in (20) comes from removing a_i.
- All off-diagonal entries of A are 1's and -1's, since the i-th equation in (20) involves all of our $2n+1$ numbers $a_1, a_2, \ldots, a_{2n+1}$ except for a_i and $a_{2n+1} = 0$.

Moreover, the vector on the right hand side of (21) is the zero vector, since the equations in (20) are homogeneous (i.e., have no constant term). Thus, (21) can be rewritten as $Av = \mathbf{0}$, where $v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{2n} \end{pmatrix} \in \mathbb{R}^{2n}$, and where $\mathbf{0}$ is the zero vector. (Here, we use $\mathbb{R}^{2n}$ to denote the vector space of all **column** vectors of size $2n$.)

Our goal is to show that all $2n$ numbers $a_1, a_2, \ldots, a_{2n}$ are 0. In other words, our goal is to show that the vector $v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{2n} \end{pmatrix}$ satisfying $Av = \mathbf{0}$ is the zero vector. In other words, our goal is to show that the nullspace of A is trivial (i.e., consists only of the zero vector). By the invertible matrix theorem, this will follow if we can show that $\det A \neq 0$.

So let us show this. Even better, we shall show that $\det A$ is an odd integer. This will automatically yield $\det A \neq 0$, since 0 is even.

The Leibniz formula for $\det A$ yields

$$\det A = \sum_{\sigma \in S_{2n}} (-1)^{\sigma} a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{2n,\sigma(2n)}, \tag{22}$$

where $a_{i,j}$ denotes the (i,j)-th entry of A. Thus, $\det A$ is an integer (since all entries of A are integers). Moreover, if we add a multiple of 2 to some entry of A, then $\det A$ also changes by a multiple of 2 (since every addend on the right hand side of (22) either stays the same or changes by a multiple of 2), and thus the parity of $\det A$ remains unchanged. Hence, on our quest to prove that $\det A$ is odd, we can freely add multiples of 2 to each entry of A.

Recall that each off-diagonal entry of A is an **odd** integer (since it is either 1 or -1). Thus, by adding appropriate multiples of 2 to these entries, we can ensure that they all become 1. Let B be the resulting $2n \times 2n$-matrix (with 0's along the diagonal and 1's in all off-diagonal positions); thus, in our example, we have

$$B = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}. \tag{23}$$

Since B is obtained from A by adding multiples of 2 to some entries, we have $\det B \equiv \det A \bmod 2$ (because if we add a multiple of 2 to some entry of A, then $\det A$ also changes by a multiple of 2). Thus, in order to prove that $\det A$ is odd, it suffices to prove that $\det B$ is odd.

But we can compute $\det B$ explicitly. Indeed, the formula (16) that we obtained in

the solution to Exercise 10.3.1 (applied to 4, 1 and -1 instead of n, a_i and b) yields

$$\det\begin{pmatrix} 1+(-1) & 1 & 1 & 1 \\ 1 & 1+(-1) & 1 & 1 \\ 1 & 1 & 1+(-1) & 1 \\ 1 & 1 & 1 & 1+(-1) \end{pmatrix}$$
$$= \underbrace{1\cdot 1\cdot \cdots \cdot 1}_{=1} + (-1)\sum_{i=1}^{4} \underbrace{1\cdot 1\cdot \cdots \cdot \widehat{1} \cdot \cdots \cdot 1}_{=1}$$
$$= 1 + (-1)\underbrace{\sum_{i=1}^{4} 1}_{=4} = 1 + (-1)\,4 = 1 - 4 = -3.$$

Since $1+(-1) = 0$, the matrix on the left hand side of this equality is precisely our B, and thus the equality rewrites as

$$\det B = -3,$$

which shows that $\det B$ is odd, and we are done. In the general case, $\det B$ will be $1+(-1)\,2n = 1-2n$, which is also odd, so the argument still works. □

See also [Grinbe20a, Exercise 5.3.3] for a different approach to Exercise 10.3.2 (stopping short of a complete solution, but solving the exercise in the case when $a_1, a_2, \ldots, a_{2n+1}$ are **rational** numbers). Two other solutions (both using linear algebra!) can be found in [GelAnd17, Problem 300].

10.3.4. Rank and nullity

Another feature of matrices that often comes useful in contest problems (and elsewhere) is the notion of rank. The *rank* of a matrix A (over any field $\mathbb{F}$) is defined to be the maximum number of linearly independent columns of A. Equivalently, it can be defined as the maximum number of linearly independent rows of A. The equivalence is nontrivial (see, e.g., [Axler23, 3.57] or [Camero08, Theorem 2.5] or [StoLui18, Theorem 8.12 (3)] for a proof) and often useful by itself. Another way to describe the rank of A is as the largest integer k such that A has a $k \times k$-submatrix with nonzero determinant (see, e.g., [Kuttle22, Corollary 8.6.8]). Yet another is as the dimension of the image of A (that is, of the vector space of all vectors of the form Av). The rank of a matrix A is commonly denoted $\operatorname{rank} A$. The most important property of the rank is probably the *rank-nullity theorem* ([StoLui18, Theorem 8.12]; see also [Axler23, 3.21] or [Griffi20, Theorem 3.40] or [StoLui18, Theorem 8.3] for a statement in terms of linear maps):

Theorem 10.3.1. Let A be an $n \times m$-matrix over a field $\mathbb{F}$. Let $\mathbb{F}^m$ denote the vector space of all column vectors of size m. Let $\operatorname{Ker} A$ be the nullspace of A

(that is, the set of all column vectors $v \in \mathbb{F}^m$ such that $Av = \mathbf{0}$, where $\mathbf{0}$ denotes the zero vector of size n). Then,

$$m = \dim(\operatorname{Ker} A) + \operatorname{rank} A.$$

This theorem can be used to compute $\operatorname{rank} A$ from $\dim(\operatorname{Ker} A)$ and vice versa, which is often helpful. We will see an example soon, but let us first explore a different topic.

10.4. Linear algebra over $\mathbb{F}_2$

10.4.1. A brief introduction to $\mathbb{F}_2$

As mentioned in Subsection 10.1 above, linear algebra can be done over any field. The most popular choices are $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$, but other options should not be discounted either.

The simplest of all fields is none of $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$. It is a field called $\mathbb{F}_2$ or $\mathbb{Z}/2$, consisting of just two elements. We call these elements $\overline{0}$ and $\overline{1}$. You can think of $\overline{0}$ as the set of all even integers, and of $\overline{1}$ as the set of all odd integers (these are their actual definitions). The addition is defined by

$$\overline{0}+\overline{0}=\overline{0}, \qquad \overline{0}+\overline{1}=\overline{1}, \qquad \overline{1}+\overline{0}=\overline{1}, \qquad \overline{1}+\overline{1}=\overline{0};$$

the multiplication is defined by

$$\overline{0}\cdot\overline{0}=\overline{0}, \qquad \overline{0}\cdot\overline{1}=\overline{0}, \qquad \overline{1}\cdot\overline{0}=\overline{0}, \qquad \overline{1}\cdot\overline{1}=\overline{1}.$$

These are just the standard parity rules for integers: For example, $\overline{1}+\overline{1}=\overline{0}$ means "odd plus odd equals even", whereas $\overline{1}\cdot\overline{0}=\overline{0}$ means "odd times even equals even".

The set $\mathbb{F}_2 = \{\overline{0}, \overline{1}\}$, equipped with the addition and the multiplication just defined, is a field. Its subtraction is precisely its addition (since plus and minus are the same modulo 2). Its division is even simpler: Division by $\overline{1}$ changes nothing, whereas division by $\overline{0}$ is impossible (you cannot divide by zero, and $\overline{0}$ is the zero of $\mathbb{F}_2$).

10.4.2. Oddtown

You might think that nothing interesting can be done with such a small field as $\mathbb{F}_2$, but you would be mistaken. Linear algebra is not about the field, but about vectors over the field; and of course, vectors over $\mathbb{F}_2$ can carry much more information than scalars in $\mathbb{F}_2$. The following exercise is a first illustration:

Exercise 10.4.1. Let $n, m \in \mathbb{N}$. In a town with n inhabitants, there are m clubs, each of which has an odd number of members. Any two distinct clubs share an even number of common members. Prove that $m \leq n$.

Solution idea. Again, we assume the contrary. Thus, $m > n$. We denote the n inhabitants by $1, 2, \ldots, n$, and we denote the m clubs by $C_1, C_2, \ldots, C_m$.

Let $v_1, v_2, \ldots, v_m$ be the indicator vectors of our m clubs over $\mathbb{F}_2$. This means that, for each $i \in \{1, 2, \ldots, m\}$, the vector v_i is a row vector of size n whose j-th entry is

$$\begin{cases} \overline{1}, & \text{if } j \in C_i; \\ \overline{0}, & \text{if } j \notin C_i \end{cases} \qquad \text{for each } j \in \{1, 2, \ldots, n\}.$$

Thus we have defined m vectors $v_1, v_2, \ldots, v_m$ in the n-dimensional vector space $\mathbb{F}_2^n$ over the field $\mathbb{F}_2$.

We shall use dot products again, as in the solution to Exercise 10.2.2 (but this time, they live in $\mathbb{F}_2$ rather than $\mathbb{R}$). Just as we proved (4) in the latter solution, we can now see that any $i, j \in \{1, 2, \ldots, m\}$ satisfy

$$\left\langle v_i, v_j \right\rangle = \overline{\left| C_i \cap C_j \right|}, \tag{24}$$

where $\overline{\left| C_i \cap C_j \right|}$ means $\begin{cases} \overline{0}, & \text{if } \left| C_i \cap C_j \right| \text{ is even;} \\ \overline{1}, & \text{if } \left| C_i \cap C_j \right| \text{ odd} \end{cases}$ (since a sum of an even number of $\overline{1}$'s is $\overline{0}$, whereas a sum of an odd number of $\overline{1}$'s is $\overline{1}$). Hence:

- For any two distinct elements i, j of $\{1, 2, \ldots, m\}$, we have

$$\left\langle v_i, v_j \right\rangle = \overline{\left| C_i \cap C_j \right|} = \overline{0} \tag{25}$$

  (since any two distinct clubs share an even number of common members).

- For any $i \in \{1, 2, \ldots, m\}$, we have

$$\left\langle v_i, v_i \right\rangle = \overline{\left| C_i \cap C_i \right|} = \overline{\left| C_i \right|} = \overline{1} \tag{26}$$

  (since each club has an odd number of members).

Our m vectors $v_1, v_2, \ldots, v_m$ belong to the n-dimensional vector space $\mathbb{F}_2^n$, and thus are linearly dependent (by Theorem 10.2.1 **(a)**, since $m > n$). In other words, there exist m elements $c_1, c_2, \ldots, c_m \in \mathbb{F}_2$, not all zero, such that

$$c_1 v_1 + c_2 v_2 + \cdots + c_m v_m = \mathbf{0}$$

(where $\mathbf{0}$ is the all-zero vector over $\mathbb{F}_2$, that is, $\left(\underbrace{\overline{0}, \overline{0}, \ldots, \overline{0}}_{n \text{ entries}} \right)$). Consider these m elements.

Since the dot product is linear in each argument, we now have

$$\begin{aligned}
&\langle c_1 v_1 + c_2 v_2 + \cdots + c_m v_m,\ v_1 \rangle \\
&= c_1 \underbrace{\langle v_1, v_1 \rangle}_{\substack{=\overline{1} \\ \text{(by (26))}}} + c_2 \underbrace{\langle v_2, v_1 \rangle}_{\substack{=\overline{0} \\ \text{(by (25))}}} + \cdots + c_m \underbrace{\langle v_m, v_1 \rangle}_{\substack{=\overline{0} \\ \text{(by (25))}}} \\
&= c_1 \overline{1} + \underbrace{c_2 \overline{0} + \cdots + c_m \overline{0}}_{=\overline{0}} = c_1 \overline{1} = c_1,
\end{aligned}$$

so that

$$c_1 = \left\langle \underbrace{c_1 v_1 + c_2 v_2 + \cdots + c_m v_m}_{=\mathbf{0}},\ v_1 \right\rangle = \langle \mathbf{0},\ v_1 \rangle = \overline{0}.$$

Similarly, we can find that $c_i = \overline{0}$ for all $i \in \{1, 2, \ldots, m\}$ (since there is nothing special about the 1-st club). But this contradicts the fact that not all $c_1, c_2, \ldots, c_m$ are zero. This contradiction shows that our assumption was wrong. Hence, Exercise 10.4.1 is solved. □

Remark 10.4.1. The inequality $m \leq n$ in Exercise 10.4.1 is sharp (i.e., equality can be achieved). In fact, if each inhabitant of our town forms a very exclusive club just by himself, then we obtain n clubs that satisfy the conditions of Exercise 10.4.1.

10.4.3. Eventown

Exercise 10.4.1 is one of the four famous *Oddtown/Eventown theorems* (which appear in many sources: e.g., three of them appear in [BabFra23, §1.1] and in [Bollob10, problems 59, 114 and 116]); here is another:

Exercise 10.4.2. Let $n, m \in \mathbb{N}$ with $n > 0$. In a town with n inhabitants, there are m clubs, each of which has an even number of members. Any two distinct clubs share an odd number of common members.

(a) Prove that $m \leq n$.

(b) If n is even, then prove that $m \leq n - 1$.

Solution idea. As in the solution to Exercise 10.4.1, we shall use row vectors of size n over the field $\mathbb{F}_2$. Again, let $v_1, v_2, \ldots, v_m$ be the indicator vectors of our m clubs over $\mathbb{F}_2$. Then, just as we proved the equalities (25) and (26) in the solution to Exercise 10.4.1, we can now show that:

- For any two distinct elements i, j of $\{1, 2, \ldots, m\}$, we have
$$\langle v_i, v_j \rangle = \overline{1} \tag{27}$$
(since any two distinct clubs share an odd number of common members).

- For any $i \in \{1, 2, \ldots, m\}$, we have

$$\langle v_i, v_i \rangle = \overline{0} \tag{28}$$

(since each club has an even number of members).

Let W be the subset

$$\begin{aligned} &\left\{ x \in \mathbb{F}_2^n \mid \text{the sum of all coordinates of } x \text{ is } \overline{0} \right\} \\ &= \left\{ (x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n \mid x_1 + x_2 + \cdots + x_n = \overline{0} \right\} \end{aligned}$$

of $\mathbb{F}_2^n$. Then, W is a vector subspace of $\mathbb{F}_2^n$ (this is easy to see[16]). It is a proper subset of $\mathbb{F}_2^n$ (since it does not contain the vector $\left(\overline{1}, \overline{0}, \overline{0}, \ldots, \overline{0}\right)$, and thus its dimension $\dim W$ is lower than the dimension of $\mathbb{F}_2^n$. In other words,

$$\dim W < \dim \left(\mathbb{F}_2^n\right) = n,$$

so that $\dim W \leq n - 1$. (Actually, $\dim W$ is exactly $n - 1$, but we don't need this fact[17].)

Let $i \in \{1, 2, \ldots, m\}$. Then, the i-th club has an even number of members (by an assumption of the exercise). In other words, the vector v_i has an even number of coordinates equal to $\overline{1}$. Hence, the sum of all coordinates of v_i is $\overline{0}$ (since the sum of an even number of $\overline{1}$'s is always $\overline{0}$). In other words, $v_i \in W$ (by the definition of W).

Forget that we fixed i. We thus have shown that $v_i \in W$ for each $i \in \{1, 2, \ldots, m\}$. In other words, the m vectors $v_1, v_2, \ldots, v_m$ belong to W.

Now we shall prove two crucial claims:

---

[16] *Proof.* The easiest way to see this is to argue that W is the left nullspace of the matrix $\begin{pmatrix} \overline{1} & \overline{1} & \cdots & \overline{1} \end{pmatrix}$, and thus is a vector subspace (like any nullspace). A more down-to-earth proof goes by arguing that if $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ are two vectors in W, then their sum $x + y = (x_1 + y_1,\ x_2 + y_2,\ \ldots,\ x_n + y_n)$ also belongs to W, since

$$\begin{aligned} &(x_1 + y_1) + (x_2 + y_2) + \cdots + (x_n + y_n) \\ &= \underbrace{(x_1 + x_2 + \cdots + x_n)}_{\substack{=\overline{0} \\ (\text{since } x \in W)}} + \underbrace{(y_1 + y_2 + \cdots + y_n)}_{\substack{=\overline{0} \\ (\text{since } y \in W)}} = \overline{0} + \overline{0} = \overline{0}. \end{aligned}$$

(We would also need to show that $\mathbf{0} \in W$ and that $\lambda x \in W$ for each $\lambda \in \mathbb{F}_2$ and each $x \in W$. But the former is trivial, and the latter is is automatic, since the only options for λ are $\overline{0}$ and $\overline{1}$.)

[17] Nevertheless, it is easy to prove: For example, one can argue that W is carved out of $\mathbb{F}_2^n$ by the single equation $x_1 + x_2 + \cdots + x_n = \overline{0}$, which is not a tautology and thus decreases the dimension by 1. Alternatively, one can easily see that W has a basis

$$\left(\overline{1}, \overline{0}, \overline{0}, \ldots, \overline{0}, \overline{1}\right),\ \left(\overline{0}, \overline{1}, \overline{0}, \overline{0}, \ldots, \overline{0}, \overline{1}\right),\ \ldots,\ \left(\overline{0}, \overline{0}, \overline{0}, \ldots, \overline{0}, \overline{1}, \overline{1}\right)$$

(each vector of this basis has a $\overline{1}$ in its last position and at another position, and $\overline{0}$'s everywhere else), and this basis has $n - 1$ elements, so that $\dim W = n - 1$.

Claim 1: If m is even, then $m \leq n-1$.

Proof of Claim 1. Assume that m is even. We must prove that $m \leq n-1$.

Assume the contrary. Thus, $m > n-1 \geq \dim W$ (since $\dim W \leq n-1$). But the m vectors $v_1, v_2, \ldots, v_m$ belong to the vector space W, whose dimension is $\dim W$. Hence, Theorem 10.2.1 **(a)** (applied to m and $\dim W$ instead of k and n) yields that these m vectors $v_1, v_2, \ldots, v_m$ are linearly dependent. In other words, there exist m elements $c_1, c_2, \ldots, c_m \in \mathbb{F}_2$, not all zero, such that

$$c_1 v_1 + c_2 v_2 + \cdots + c_m v_m = \mathbf{0}$$

(where $\mathbf{0}$ is the all-zero vector over $\mathbb{F}_2$, that is, $\left(\underbrace{\overline{0}, \overline{0}, \ldots, \overline{0}}_{n \text{ entries}}\right)$). Consider these m elements.

Set $d := c_1 + c_2 + \cdots + c_m$.

Since the dot product is linear in each argument, we now have

$$\begin{aligned}
&\langle c_1 v_1 + c_2 v_2 + \cdots + c_m v_m,\ v_1 \rangle \\
&= c_1 \underbrace{\langle v_1, v_1 \rangle}_{\substack{=\overline{0} \\ \text{(by (28))}}} + c_2 \underbrace{\langle v_2, v_1 \rangle}_{\substack{=\overline{1} \\ \text{(by (27))}}} + \cdots + c_m \underbrace{\langle v_m, v_1 \rangle}_{\substack{=\overline{1} \\ \text{(by (27))}}} \\
&= \underbrace{c_1 \overline{0}}_{=\overline{0}} + \underbrace{c_2 \overline{1}}_{=c_2} + \cdots + \underbrace{c_m \overline{1}}_{=c_m} = \overline{0} + c_2 + \cdots + c_m \\
&= c_2 + c_3 + \cdots + c_m = \underbrace{(c_1 + c_2 + \cdots + c_m)}_{=d} - c_1 = d - c_1,
\end{aligned}$$

so that

$$d - c_1 = \left\langle \underbrace{c_1 v_1 + c_2 v_2 + \cdots + c_m v_m}_{=\mathbf{0}},\ v_1 \right\rangle = \langle \mathbf{0},\ v_1 \rangle = \overline{0}.$$

Hence, $d = c_1$, so that $c_1 = d$. Similarly, we can find that $c_i = d$ for all $i \in \{1, 2, \ldots, m\}$ (since there is nothing special about the 1-st club). In other words, all the elements $c_1, c_2, \ldots, c_m$ equal d. Hence, d is not zero (since not all $c_1, c_2, \ldots, c_m$ are zero). In other words, $d \neq \overline{0}$. Since $d \in \mathbb{F}_2 = \{\overline{0}, \overline{1}\}$, we thus obtain $d = \overline{1}$.

Thus, for all $i \in \{1, 2, \ldots, m\}$, we have $c_i = d = \overline{1}$. Adding these equations together for all $i \in \{1, 2, \ldots, m\}$, we obtain

$$c_1 + c_2 + \cdots + c_m = \underbrace{\overline{1} + \overline{1} + \cdots + \overline{1}}_{m \text{ many } \overline{1}\text{'s}} = \overline{0}$$

(because m is even, but a sum of an even number of $\overline{1}$'s is always $\overline{0}$). This contradicts $c_1 + c_2 + \cdots + c_m = d = \overline{1}$. This contradiction shows that our assumption was wrong. Hence, Claim 1 is proved. $\square$

Claim 2: If m is odd, then $m \leq n$.

Proof of Claim 2. Assume that m is odd. Thus, $m-1$ is even. Furthermore, m cannot be 0 (since m is odd), and thus $m-1 \in \mathbb{N}$.

Now, let us dissolve the m-th club. Then, we are left with $m-1$ clubs, which still satisfy all assumptions of the exercise (i.e., each of them has an even number of members, but any two have an odd number of common members). Thus, we can apply Claim 1 to these $m-1$ clubs (using $m-1$ instead of m), and conclude that $m-1 \leq n-1$ (since $m-1$ is even). Thus, $m \leq n$. This proves Claim 2. □

Now we can easily solve the exercise:

(a) We must prove that $m \leq n$. If m is odd, then this follows directly from Claim 2. If m is even, then this follows from Claim 1, since Claim 1 yields $m \leq n-1 \leq n$. In either case, Exercise 10.4.2 **(a)** is solved.

(b) Assume that n is even. We must prove that $m \leq n-1$.

Assume the contrary. Thus, $m > n-1$, so that $m \geq n$. But part **(a)** yields $m \leq n$. Combining these two inequalities, we find $m = n$, so that m is even (since n is even). Hence, Claim 1 yields $m \leq n-1$, which contradicts $m \geq n > n-1$. This contradiction shows that our assumption was false. Exercise 10.4.2 **(b)** is thus solved. □

Remark 10.4.2. When n is even, the inequality $m \leq n-1$ in Exercise 10.4.2 **(b)** is sharp (i.e., equality can be achieved). In fact, let us single out an inhabitant α. For each of the $n-1$ remaining inhabitants $\beta \neq \alpha$, we form a club containing only α and β. Thus, we obtain $n-1$ clubs that satisfy the conditions of Exercise 10.4.1.

When n is odd, the inequality $m \leq n$ in Exercise 10.4.2 **(a)** is sharp as well. To see it become an equality, we construct $n-1$ clubs just as in the preceding paragraph, but we also create an n-th club that contains all $n-1$ inhabitants distinct from α. The resulting n clubs satisfy the conditions of Exercise 10.4.1 (since n is odd).

Remark 10.4.3. There is an alternative solution to Exercise 10.4.2, which is worth sketching since it connects it to our previous work on determinants.

Again, we let $v_1, v_2, \ldots, v_m$ be the indicator vectors of our m clubs over $\mathbb{F}_2$. This time, we pack them into a matrix: Let A be the $n \times m$-matrix with columns $v_1, v_2, \ldots, v_m$ over $\mathbb{F}_2$. Consider also the transpose A^T of this matrix A. Then, the equalities (28) and (27) show that $A^T A = R$, where R is the $m \times m$-matrix

$$\begin{pmatrix} \overline{0} & \overline{1} & \overline{1} & \cdots & \overline{1} \\ \overline{1} & \overline{0} & \overline{1} & \cdots & \overline{1} \\ \overline{1} & \overline{1} & \overline{0} & \cdots & \overline{1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{1} & \overline{1} & \overline{1} & \cdots & \overline{0} \end{pmatrix}$$

whose all diagonal entries equal $\overline{0}$ and whose all off-diagonal entries equal $\overline{1}$. Again, defining W to be the subset

$$\left\{x \in \mathbb{F}_2^n \mid \text{the sum of all coordinates of } x \text{ is } \overline{0}\right\}$$

of $\mathbb{F}_2^n$, we can easily see that W is a vector subspace of dimension $\leq n-1$ (actually $= n-1$), and therefore

$$\begin{aligned} n-1 &\geq \dim W \geq \operatorname{rank} A \qquad \text{(since all columns of } A \text{ belong to } W\text{)} \\ &\geq \operatorname{rank}\left(A^T A\right) \qquad \left(\begin{array}{c}\text{since } \operatorname{rank} V \geq \operatorname{rank}(UV) \text{ for} \\ \text{any two matrices } U \text{ and } V\end{array}\right) \\ &= \operatorname{rank} R \qquad \left(\text{since } A^T A = R\right). \end{aligned}$$

It remains to compute $\operatorname{rank} R$. Here, it helps to observe that R is the $\mathbb{F}_2$-analogue of the matrix B considered in our solution to Exercise 10.3.2 above (see (23)), and its determinant can be computed along the same lines (or it can be obtained from $\det B$ by "projecting" onto $\mathbb{F}_2$). This shows that $\det R = \overline{1}$ if m is even. When m is odd, we have $\det R = \overline{0}$, but the $(m-1) \times (m-1)$-submatrix of R obtained by removing the last row and the last column has determinant $\overline{1}$. Combining these facts, we conclude that

$$\operatorname{rank} R = \begin{cases} m, & \text{if } m \text{ is even;} \\ m-1, & \text{if } m \text{ is odd} \end{cases}$$

(since the rank of a matrix M equals the largest $k \in \mathbb{N}$ such that M has a $k \times k$-submatrix with nonzero determinant). Combining what we have shown, we find

$$n-1 \geq \operatorname{rank} R = \begin{cases} m, & \text{if } m \text{ is even;} \\ m-1, & \text{if } m \text{ is odd,} \end{cases}$$

which yields both Claims 1 and 2 from our above solution to Exercise 10.4.2.

10.4.4. Size and dimension

There are two more Eventown/Oddtown theorems: see Exercises 10.5.2 and 10.5.3 below. In preparation for their proofs, let us prove a simple property of finite-dimensional vector spaces over $\mathbb{F}_2$. In "usual" linear algebra (i.e., linear algebra over $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$), you rarely have a reason to think about the size of a vector space, since all nontrivial vector spaces are infinite. However, when working over a finite field such as $\mathbb{F}_2$, finite-dimensional vector spaces are finite sets, and it makes sense to ask about their sizes. The answer (which we only state here for $\mathbb{F}_2$) is nice and simple:

Proposition 10.4.4. Let $n \in \mathbb{N}$. Let V be a vector space over $\mathbb{F}_2$ that has dimension n. Then, the size of V (as a set) is $|V| = 2^n$.

Proof. The vector space V has dimension n; thus, it has a basis $(v_1, v_2, \ldots, v_n)$ consisting of n vectors. Consider this basis.

Each vector $w \in V$ can be uniquely expressed as a linear combination $c_1v_1 + c_2v_2 + \cdots + c_nv_n$ with coefficients $c_1, c_2, \ldots, c_n \in \mathbb{F}_2$ (since $(v_1, v_2, \ldots, v_n)$ is a basis of V). In other words, the map

$$\begin{aligned} \mathbb{F}_2^n &\to V, \\ (c_1, c_2, \ldots, c_n) &\mapsto c_1v_1 + c_2v_2 + \cdots + c_nv_n \end{aligned}$$

is a bijection[18]. By the bijection principle[19], this entails $|\mathbb{F}_2^n| = |V|$. Hence, $|V| = |\mathbb{F}_2^n| = |\mathbb{F}_2|^n = 2^n$ (since $|\mathbb{F}_2| = 2$). This proves Proposition 10.4.4. □

10.5. Orthogonality

10.5.1. Theory

Let us now discuss a bit more linear algebra, which will help us prove two more Eventown/Oddtown theorems.

Let $\mathbb{F}$ be any field (for example, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{F}_2$), and let $n \in \mathbb{N}$ be arbitrary. Consider the vector space $\mathbb{F}^n$ of all row vectors $(a_1, a_2, \ldots, a_n)$ of size n with entries in $\mathbb{F}$. The *dot product* of two vectors $x = (x_1, x_2, \ldots, x_n) \in \mathbb{F}^n$ and $y = (y_1, y_2, \ldots, y_n) \in \mathbb{F}^n$ is defined to be the scalar

$$x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{p=1}^{n} x_py_p \in \mathbb{F}.$$

This dot product is denoted by $\langle x, y \rangle$ or sometimes by $x \cdot y$. (We have already used this dot product for $\mathbb{F} = \mathbb{R}$ in the solution to Exercise 10.2.2, and for $\mathbb{F} = \mathbb{F}_2$ in Subsection 10.4. Now we are defining it in the general case.)

The dot product has several important properties (all of which follow easily from its definition):

- *Symmetry*: We have $\langle x, y \rangle = \langle y, x \rangle$ for any $x, y \in \mathbb{F}^n$.

[18] Indeed,

- it is surjective since each vector $w \in V$ can be expressed as a linear combination $c_1v_1 + c_2v_2 + \cdots + c_nv_n$ with coefficients $c_1, c_2, \ldots, c_n \in \mathbb{F}_2$;
- and it is injective since these coefficients $c_1, c_2, \ldots, c_n \in \mathbb{F}_2$ are uniquely determined by w.

[19] The *bijection principle* says that if there exists a bijection between two sets X and Y, then $|X| = |Y|$. This is a fundamental fact that underlies all enumerative combinatorics.

- *Bilinearity:* We have

$$\begin{aligned}
\langle x + x',\ y\rangle &= \langle x, y\rangle + \langle x', y\rangle && \text{for all } x, x', y \in \mathbb{F}^n;\\
\langle x,\ y + y'\rangle &= \langle x, y\rangle + \langle x, y'\rangle && \text{for all } x, y, y' \in \mathbb{F}^n;\\
\langle \lambda x, y\rangle &= \langle x, \lambda y\rangle = \lambda \langle x, y\rangle && \text{for all } \lambda \in \mathbb{F} \text{ and } x, y \in \mathbb{F}^n;\\
\langle \mathbf{0}, x\rangle &= \langle x, \mathbf{0}\rangle = 0 && \text{for all } x \in \mathbb{F}^n,
\end{aligned}$$

  where $\mathbf{0}$ denotes the zero vector $(0, 0, \ldots, 0)$ in $\mathbb{F}^n$.

For two vectors x and y in $\mathbb{F}^n$, we say that x is *orthogonal* to y if and only if $\langle x, y\rangle = 0$. The shorthand notation for this is "$x \perp y$". This orthogonality relation is symmetric (i.e., we have $x \perp y$ if and only if $y \perp x$). If $\mathbb{F} = \mathbb{R}$, then it is precisely the classical orthogonality relation from Euclidean geometry (i.e., two vectors are orthogonal if and only if they span perpendicular lines or one of them is zero). But it is useful for other fields as well, even $\mathbb{F}_2$! (Of course, when $\mathbb{F} = \mathbb{F}_2$, the relation $\langle x, y\rangle = 0$ means $\langle x, y\rangle = \overline{0}$, since $\overline{0}$ plays the role of 0 in $\mathbb{F}_2$.)

Warning 10.5.1. Not all properties known from the real case ($\mathbb{F} = \mathbb{R}$) generalize to other fields! For $\mathbb{F} = \mathbb{R}$, the only vector $x \in \mathbb{F}^n$ satisfying $x \perp x$ is the zero vector $\mathbf{0}$. But this is not the case for $\mathbb{F} = \mathbb{F}_2$. For instance, the vector $\left(\overline{1}, \overline{1}\right) \in \mathbb{F}_2^2$ is orthogonal to itself, since $\left\langle \left(\overline{1}, \overline{1}\right),\ \left(\overline{1}, \overline{1}\right)\right\rangle = \overline{1}\cdot\overline{1} + \overline{1}\cdot\overline{1} = \overline{1} + \overline{1} = \overline{0}$. But it is not the zero vector. A similar example exists for $\mathbb{F} = \mathbb{C}$: Here, the vector $(1, i)$ (with $i = \sqrt{-1}$) is orthogonal to itself but is not zero. The reason why such vectors don't exist for $\mathbb{F} = \mathbb{R}$ is that in $\mathbb{R}$, a sum of squares is never zero unless all addends are zero.

This, incidentally, is the reason why many authors shun the dot product for $\mathbb{F} = \mathbb{C}$ in favor of a subtler notion of product, which involves complex conjugation (specifically, it replaces $x_1y_1 + x_2y_2 + \cdots + x_ny_n$ by $x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}$ in the definition of the inner product); this latter product is usually called the *inner product*. See [Axler23, Chapter 6] for much more about it.

The following fact is easy but crucial:

Lemma 10.5.2. Let $n \in \mathbb{N}$, and let $\mathbb{F}$ be a field. Let $u_1, u_2, \ldots, u_n$ and v be some vectors in $\mathbb{F}^n$. Assume that v is orthogonal to each of the vectors $u_1, u_2, \ldots, u_n$. Then, v is also orthogonal to each linear combination of these vectors $u_1, u_2, \ldots, u_n$.

Proof sketch. Let w be a linear combination of $u_1, u_2, \ldots, u_n$. Thus, $w = \lambda_1u_1 + \lambda_2u_2 + \cdots + \lambda_nu_n$ for some coefficients $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$. Consider these $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Thus,

$$\begin{aligned}\langle v, w\rangle &= \langle v,\ \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n\rangle \\ &= \lambda_1 \langle v, u_1\rangle + \lambda_2 \langle v, u_2\rangle + \cdots + \lambda_n \langle v, u_n\rangle \\ &\qquad \text{(by the bilinearity of the dot product)} \\ &= \sum_{k=1}^{n} \lambda_k \underbrace{\langle v, u_k\rangle}_{\substack{=0 \\ \text{(since } v \text{ is orthogonal} \\ \text{to each of } u_1, u_2, \ldots, u_n, \\ \text{thus in particular to } u_k\text{)}}} = \sum_{k=1}^{n} \lambda_k 0 = 0.\end{aligned}$$

In other words, v is orthogonal to w. Since w was chosen to be an arbitrary linear combination of $u_1, u_2, \ldots, u_n$, we thus have shown that v is orthogonal to each such combination. This proves Lemma 10.5.2. □

An important definition related to orthogonality is the "orthogonal space". If U is any subset of $\mathbb{F}^n$ (for some $n \in \mathbb{N}$ and some field $\mathbb{F}$), then we define a subset $U^\perp$ of $\mathbb{F}^n$ by

$$U^\perp := \left\{ v \in \mathbb{F}^n \ \mid\ v \perp u \text{ for each } u \in U \right\}.$$

This is the set of all vectors $v \in \mathbb{F}^n$ that are orthogonal to **all** vectors in U. It is called the *orthogonal complement* of U in $\mathbb{F}^n$. (Informally, it is often called the "*perp* of U" or the "U-perp", imitating the LaTeX command `\perp` for the $\perp$ symbol.) Its major properties are the following:[20]

Theorem 10.5.3. Let $n \in \mathbb{N}$, and let $\mathbb{F}$ be a field. Let U be a subset of $\mathbb{F}^n$. Then:

(a) The set $U^\perp$ is a vector subspace of $\mathbb{F}^n$ (even if U is not!).

(b) We have $U \subseteq \left(U^\perp\right)^\perp$.

(c) Assume that U is a vector subspace of $\mathbb{F}^n$. Then,

$$\dim U + \dim\left(U^\perp\right) = n.$$

(d) Assume that U is a vector subspace of $\mathbb{F}^n$. Then, $\left(U^\perp\right)^\perp = U$.

Proof sketch. Part **(a)** follows from the bilinearity of the dot product (or from Lemma 10.5.2); part **(b)** from its symmetry.

(c) This is the rank-nullity theorem in disguise. To wit: Let $(u_1, u_2, \ldots, u_k)$ be a basis of U. Thus, $\dim U = k$. Note that the k vectors $u_1, u_2, \ldots, u_k$ are linearly

[20] Theorem 10.5.3 can be generalized by replacing $\mathbb{F}^n$ with an arbitrary finite-dimensional vector space, and replacing the dot product with an arbitrary bilinear form. See, e.g., [Grinbe20b, Corollary 7.1] for this kind of generalization of Theorem 10.5.3 **(c)**.

independent (since they form a basis of U), and their span is $\operatorname{span}\{u_1, u_2, \ldots, u_k\} = U$ (for the same reason).

Now let A be the $k \times n$-matrix (with entries in $\mathbb{F}$) whose rows are these k vectors $u_1, u_2, \ldots, u_k$. Then, the k rows of A are linearly independent (since we just showed that $u_1, u_2, \ldots, u_k$ are linearly independent), and thus the rank of A is $\operatorname{rank} A = k$.

Next, we consider the nullspace $\operatorname{Ker} A$ of A. This is defined as the set of all column vectors $v \in \mathbb{F}^n$ such that $Av = \mathbf{0}$. In other words,

$$\operatorname{Ker} A = \{v \in \mathbb{F}^n \mid Av = \mathbf{0}\}. \tag{29}$$

Being sloppy, we regard the column vectors in $\operatorname{Ker} A$ as row vectors (by identifying them with their transposes).

The rows of A are $u_1, u_2, \ldots, u_k$. Thus, for any column vector $v \in \mathbb{F}^n$, the product Av is just the vector $\begin{pmatrix} \langle u_1, v\rangle \\ \langle u_2, v\rangle \\ \vdots \\ \langle u_k, v\rangle \end{pmatrix}$ (why?). Hence, for any $v \in \mathbb{F}^n$, we have the following chain of logical equivalences:

$$\begin{aligned}
&(Av = \mathbf{0}) \\
&\iff \left(\begin{pmatrix} \langle u_1, v\rangle \\ \langle u_2, v\rangle \\ \vdots \\ \langle u_k, v\rangle \end{pmatrix} = \mathbf{0} \right) \\
&\iff (\langle u_i, v\rangle = 0 \text{ for each } i \in \{1, 2, \ldots, k\}) \\
&\iff (\langle u, v\rangle = 0 \text{ for each } u \in \operatorname{span}\{u_1, u_2, \ldots, u_k\}) \\
&\qquad \left(\begin{array}{c} \text{since each } u \in \operatorname{span}\{u_1, u_2, \ldots, u_k\} \text{ can be written as} \\ \text{a linear combination } c_1u_1 + c_2u_2 + \cdots + c_ku_k \text{ with} \\ \text{coefficients } c_1, c_2, \ldots, c_k \in \mathbb{F}, \text{ and thus} \\ \text{satisfies } \langle u, v\rangle = \langle c_1u_1 + c_2u_2 + \cdots + c_ku_k,\ v\rangle \\ = c_1 \langle u_1, v\rangle + c_2 \langle u_2, v\rangle + \cdots + c_k \langle u_k, v\rangle \\ \text{(by the bilinearity of the dot product); but this latter} \\ \text{sum is 0 if we have } \langle u_i, v\rangle = 0 \text{ for each } i \in \{1, 2, \ldots, k\} \end{array} \right) \\
&\iff (\langle u, v\rangle = 0 \text{ for each } u \in U) \qquad (\text{since } \operatorname{span}\{u_1, u_2, \ldots, u_k\} = U) \\
&\iff (\langle v, u\rangle = 0 \text{ for each } u \in U) \qquad (\text{since } \langle u, v\rangle = \langle v, u\rangle \text{ for any } u \in \mathbb{F}^n) \\
&\iff (v \perp u \text{ for each } u \in U) \qquad (\text{since } v \perp u \text{ means that } \langle v, u\rangle = 0).
\end{aligned}$$

Hence, we can rewrite (29) as

$$\begin{aligned}
\operatorname{Ker} A &= \{v \in \mathbb{F}^n \mid v \perp u \text{ for each } u \in U\} \\
&= U^\perp \qquad \left(\text{by the definition of } U^\perp\right).
\end{aligned}$$

But the rank-nullity theorem (Theorem 10.3.1, applied to k and n instead of n and m) yields

$$n = \dim\left(\underbrace{\operatorname{Ker} A}_{=U^\perp}\right) + \underbrace{\operatorname{rank} A}_{=k=\dim U} = \dim\left(U^\perp\right) + \dim U = \dim U + \dim\left(U^\perp\right).$$

This proves Theorem 10.5.3 **(c)**.

(d) Theorem 10.5.3 **(a)** yields that $U^\perp$ is a vector subspace of $\mathbb{F}^n$. Theorem 10.5.3 **(a)** (applied to $U^\perp$ instead of U) yields that $\left(U^\perp\right)^\perp$ is a vector subspace of $\mathbb{F}^n$ as well. Theorem 10.5.3 **(c)** yields $\dim U + \dim\left(U^\perp\right) = n$. Theorem 10.5.3 **(c)** (applied to $U^\perp$ instead of U) yields $\dim\left(U^\perp\right) + \dim\left(\left(U^\perp\right)^\perp\right) = n$. Comparing these two equalities, we find $\dim U + \dim\left(U^\perp\right) = \dim\left(U^\perp\right) + \dim\left(\left(U^\perp\right)^\perp\right)$. Subtracting $\dim\left(U^\perp\right)$ from both sides of this, we obtain $\dim U = \dim\left(\left(U^\perp\right)^\perp\right)$.

However, Theorem 10.5.3 **(b)** yields $U \subseteq \left(U^\perp\right)^\perp$. Thus, U is a vector subspace of $\left(U^\perp\right)^\perp$. As we know, it satisfies $\dim U = \dim\left(\left(U^\perp\right)^\perp\right)$. Hence, Theorem 10.2.7 **(b)** (applied to $\left(U^\perp\right)^\perp$ and U instead of V and W) yields $\left(U^\perp\right)^\perp = U$. This proves Theorem 10.5.3 **(d)**. □

10.5.2. Application: Odd intersections

The best-known applications of orthogonal complements are found in geometry (where they are used to construct, e.g., the normal vector to a plane in space) and in representation theory (where they can be used to decompose larger representations into smaller ones). But they also have uses in combinatorics. The following example is another result in the vein of Oddtown and Eventown (although we don't state it in terms of towns and clubs):

Exercise 10.5.1. Let $n \in \mathbb{N}$, and let p and q be two positive integers.

Let S be an n-element set. Let $A_1, A_2, \ldots, A_p$ be p distinct subsets of S. Let $B_1, B_2, \ldots, B_q$ be q distinct subsets of S. (Note that "distinct" does not mean "disjoint"!)

Assume that $\left|A_i \cap B_j\right|$ is odd for every $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$. Then:

(a) There exist integers $\alpha \in \mathbb{N}$ and $\beta \in \mathbb{N}$ such that $\alpha + \beta \leq n - 1$ and $p \leq 2^\alpha$ and $q \leq 2^\beta$.

(b) In particular, we have $pq \leq 2^{n-1}$.

Solution idea. We WLOG assume that $S = \{1, 2, \ldots, n\}$ (otherwise, we rename the elements of S).

For each $i \in \{1,2,\ldots,p\}$, we let a_i denote the indicator vector of the subset A_i over $\mathbb{F}_2$; this is the vector of size n with entries in $\mathbb{F}_2$ whose k-th coordinate is

$$\begin{cases} \overline{1}, & \text{if } k \in A_i; \\ \overline{0}, & \text{if } k \notin A_i \end{cases} \qquad \text{for each } k \in \{1,2,\ldots,n\}.$$

Likewise, for each $j \in \{1,2,\ldots,q\}$, we let b_j denote the indicator vector of the subset B_j of $\mathbb{F}_2$.

The subsets $A_1, A_2, \ldots, A_p$ are distinct; thus, their indicator vectors $a_1, a_2, \ldots, a_p$ are distinct as well. Similarly, the vectors $b_1, b_2, \ldots, b_q$ are distinct.

Recall that $\left|A_i \cap B_j\right|$ is odd for every $i \in \{1,2,\ldots,p\}$ and $j \in \{1,2,\ldots,q\}$. Translating this into the language of dot products, we obtain the following:

Claim 1: We have $\left\langle a_i, b_j \right\rangle = \overline{1}$ for every $i \in \{1,2,\ldots,p\}$ and $j \in \{1,2,\ldots,q\}$.

Proof of Claim 1. Let $i \in \{1,2,\ldots,p\}$ and $j \in \{1,2,\ldots,q\}$. Then, $\left|A_i \cap B_j\right|$ is odd (by the assumption of the exercise). However, just as we proved (24) in the solution to Exercise 10.4.1, we can show that $\left\langle a_i, b_j \right\rangle = \overline{\left|A_i \cap B_j\right|}$. But $\overline{\left|A_i \cap B_j\right|} = \overline{1}$, since $\left|A_i \cap B_j\right|$ is odd. Thus, $\left\langle a_i, b_j \right\rangle = \overline{\left|A_i \cap B_j\right|} = \overline{1}$. This proves Claim 1. $\square$

Let

$$\mathbf{A} := \operatorname{span}\left\{a_1, a_2, \ldots, a_p\right\} \qquad \text{and} \qquad \mathbf{B} := \operatorname{span}\left\{b_1, b_2, \ldots, b_q\right\}.$$

Then, both $\mathbf{A}$ and $\mathbf{B}$ are subspaces of the vector space $\mathbb{F}_2^n$ (since a span of a set of vectors is always a subspace). Obviously, $a_1, a_2, \ldots, a_p \in \mathbf{A}$ and $b_1, b_2, \ldots, b_q \in \mathbf{B}$.

Let furthermore

$$\mathbf{A}' := \operatorname{span} \underbrace{\left\{a_i - a_1 \mid i \in \{1,2,\ldots,p\}\right\}}_{=\left\{a_1-a_1,\ a_2-a_1,\ \ldots,\ a_p-a_1\right\}} \qquad \text{and}$$

$$\mathbf{B}' := \operatorname{span} \underbrace{\left\{b_j - b_1 \mid j \in \{1,2,\ldots,q\}\right\}}_{=\left\{b_1-b_1,\ b_2-b_1,\ \ldots,\ b_q-b_1\right\}}.$$

Again, $\mathbf{A}'$ and $\mathbf{B}'$ are subspaces of $\mathbb{F}_2^n$. Furthermore, Claim 1 easily entails the following:

Claim 2: We have

$$\mathbf{A}' \subseteq \mathbf{B}^{\perp}. \tag{30}$$

Proof of Claim 2. Let $i \in \{1,2,\ldots,p\}$. For each $j \in \{1,2,\ldots,q\}$, we have $a_i - a_1 \perp b_j$, since

$$\left\langle a_i - a_1,\ b_j \right\rangle = \underbrace{\left\langle a_i, b_j \right\rangle}_{\substack{=\overline{1} \\ \text{(by Claim 1)}}} - \underbrace{\left\langle a_1, b_j \right\rangle}_{\substack{=\overline{1} \\ \text{(by Claim 1)}}} \qquad \left(\begin{array}{c}\text{by the bilinearity of} \\ \text{the dot product}\end{array}\right)$$

$$= \overline{1} - \overline{1} = \overline{0}.$$

In other words, the vector $a_i - a_1$ is orthogonal to each of the q vectors $b_1, b_2, \ldots, b_q$. Hence, $a_i - a_1$ is also orthogonal to each linear combination of these q vectors $b_1, b_2, \ldots, b_q$ (by Lemma 10.5.2). In other words, $a_i - a_1$ is orthogonal to each vector in $\operatorname{span}\left(b_1, b_2, \ldots, b_q\right)$. In other words, $a_i - a_1$ is orthogonal to each vector in $\mathbf{B}$ (since $\mathbf{B} = \operatorname{span}\left(b_1, b_2, \ldots, b_q\right)$). In other words, $a_i - a_1 \in \mathbf{B}^\perp$ (by the definition of $\mathbf{B}^\perp$).

Now forget that we fixed i. We thus have shown that $a_i - a_1 \in \mathbf{B}^\perp$ for each $i \in \{1, 2, \ldots, p\}$. In other words, the p vectors $a_i - a_1$ for all $i \in \{1, 2, \ldots, p\}$ all belong to $\mathbf{B}^\perp$. Thus, any linear combination of these p vectors also belongs to $\mathbf{B}^\perp$ (since $\mathbf{B}^\perp$ is a subspace of $\mathbb{F}_2^n$ and thus is closed under linear combination). In other words, $\operatorname{span}\left\{a_i - a_1 \mid i \in \{1, 2, \ldots, p\}\right\} \subseteq \mathbf{B}^\perp$. In other words, $\mathbf{A}' \subseteq \mathbf{B}^\perp$ (since $\mathbf{A}' = \operatorname{span}\left\{a_i - a_1 \mid i \in \{1, 2, \ldots, p\}\right\}$). This proves Claim 2. $\square$

Moreover, it is not hard to see the following:

Claim 3: We have

$$\dim\left(\mathbf{A}'\right) = \dim \mathbf{A} - 1. \tag{31}$$

Proof of Claim 3. Let us first observe that a_1 is not orthogonal to b_1 (since Claim 1 yields $\langle a_1, b_1 \rangle = \overline{1} \neq \overline{0}$). Thus, $a_1 \notin \mathbf{B}^\perp$ (since $b_1 \in \mathbf{B}$). Hence, $a_1 \notin \mathbf{A}'$ (since (30) shows that $\mathbf{A}' \subseteq \mathbf{B}^\perp$). But obviously, $a_1 \in \mathbf{A}$. Hence, $\mathbf{A}' \neq \mathbf{A}$ (because if we had $\mathbf{A}' = \mathbf{A}$, then we would have $a_1 \notin \mathbf{A}' = \mathbf{A}$, which would contradict $a_1 \in \mathbf{A}$).

The vectors $a_i - a_1$ for all $i \in \{1, 2, \ldots, p\}$ obviously are linear combinations of $a_1, a_2, \ldots, a_p$, and thus belong to $\operatorname{span}\left(a_1, a_2, \ldots, a_p\right) = \mathbf{A}$. Hence, all their linear combinations also belong to $\mathbf{A}$ (since $\mathbf{A}$ is a subspace of $\mathbb{F}_2^n$, and thus is closed under linear combination). In other words, $\mathbf{A}' \subseteq \mathbf{A}$ (since $\mathbf{A}' = \operatorname{span}\left\{a_i - a_1 \mid i \in \{1, 2, \ldots, p\}\right\}$ is the set of the linear combinations of the vectors $a_i - a_1$). Thus, $\mathbf{A}'$ is a subspace of $\mathbf{A}$. Consequently, Theorem 10.2.7 **(a)** yields $\dim\left(\mathbf{A}'\right) \leq \dim \mathbf{A}$.

Furthermore, if we had $\dim\left(\mathbf{A}'\right) = \dim \mathbf{A}$, then Theorem 10.2.7 **(b)** would yield $\mathbf{A} = \mathbf{A}'$ (since $\mathbf{A}'$ is a subspace of $\mathbf{A}$), which would contradict $\mathbf{A}' \neq \mathbf{A}$. Hence, $\dim\left(\mathbf{A}'\right) \neq \dim \mathbf{A}$. Combined with $\dim\left(\mathbf{A}'\right) \leq \dim \mathbf{A}$, this yields $\dim\left(\mathbf{A}'\right) < \dim \mathbf{A}$, so that $\dim\left(\mathbf{A}'\right) \leq \dim \mathbf{A} - 1$.

Now, let us show that $\dim\left(\mathbf{A}'\right) \geq \dim \mathbf{A} - 1$ as well. Indeed, choose any basis $\left(u_1, u_2, \ldots, u_k\right)$ of $\mathbf{A}'$; thus, $\dim\left(\mathbf{A}'\right) = k$. Then, the $k+1$ vectors $u_1, u_2, \ldots, u_k, a_1$ span the vector space $\mathbf{A}$ (this is easy to see[21]). According to Theorem 10.2.1 **(c)**

[21] *Proof.* Let $i \in \{1, 2, \ldots, p\}$. Then, $a_i - a_1 \in \mathbf{A}'$ (since $\mathbf{A}'$ was defined as the span of p vectors, one of which is $a_i - a_1$). Thus, $a_i - a_1 \in \mathbf{A}' = \operatorname{span}\left\{u_1, u_2, \ldots, u_k\right\}$ (since $\left(u_1, u_2, \ldots, u_k\right)$ is a basis of $\mathbf{A}'$). In other words, $a_i - a_1$ is a linear combination of $u_1, u_2, \ldots, u_k$. That is, there exist coefficients $c_1, c_2, \ldots, c_k \in \mathbb{F}_2$ satisfying $a_i - a_1 = c_1 u_1 + c_2 u_2 + \cdots + c_k u_k$. Consider these $c_1, c_2, \ldots, c_k$. Solving $a_i - a_1 = c_1 u_1 + c_2 u_2 + \cdots + c_k u_k$ for a_i, we find

$$a_i = c_1 u_1 + c_2 u_2 + \cdots + c_k u_k + a_1 \in \operatorname{span}\left\{u_1, u_2, \ldots, u_k, a_1\right\}.$$

Forget that we fixed i. We thus have shown that $a_i \in \operatorname{span}\left\{u_1, u_2, \ldots, u_k, a_1\right\}$ for each $i \in \{1, 2, \ldots, p\}$. In other words, each of the p vectors $a_1, a_2, \ldots, a_p$ belongs to $\operatorname{span}\left\{u_1, u_2, \ldots, u_k, a_1\right\}$. Hence, any linear combination of these p vectors also belongs to

(applied to $k+1$ and $\dim \mathbf{A}$ instead of k and n), this would be impossible if $k+1$ was smaller than $\dim \mathbf{A}$. Thus, $k+1 \geq \dim \mathbf{A}$. Hence, $k \geq \dim \mathbf{A} - 1$. Since $\dim(\mathbf{A}') = k$, this rewrites as $\dim(\mathbf{A}') \geq \dim \mathbf{A} - 1$. Combining this with $\dim(\mathbf{A}') \leq \dim \mathbf{A} - 1$, we obtain $\dim(\mathbf{A}') = \dim \mathbf{A} - 1$. This proves Claim 3. □

Now, let us set $\alpha := \dim(\mathbf{A}')$ and $\beta := \dim(\mathbf{B}')$. Thus, we can rewrite (31) as $\alpha = \dim \mathbf{A} - 1$. Hence, $\dim \mathbf{A} = \alpha + 1$. Similarly, $\dim \mathbf{B} = \beta + 1$. But Claim 2 yields $\mathbf{A}' \subseteq \mathbf{B}^\perp$, so that $\mathbf{A}'$ is a subspace of $\mathbf{B}^\perp$. Thus, by Theorem 10.2.7 **(a)**, we obtain $\dim(\mathbf{A}') \leq \dim\left(\mathbf{B}^\perp\right)$. However, Theorem 10.5.3 **(c)** (applied to $\mathbb{F} = \mathbb{F}_2$ and $U = \mathbf{B}$) yields

$$\dim \mathbf{B} + \dim\left(\mathbf{B}^\perp\right) = n.$$

Hence, $\dim\left(\mathbf{B}^\perp\right) = n - \underbrace{\dim \mathbf{B}}_{=\beta+1} = n - (\beta+1)$. Thus,

$$\alpha = \dim(\mathbf{A}') \leq \dim\left(\mathbf{B}^\perp\right) = n - (\beta+1),$$

so that $\alpha + \beta + 1 \leq n$. In other words, $\alpha + \beta \leq n - 1$.

Next, we recall that the p vectors $a_1, a_2, \ldots, a_p$ are distinct. In other words, the p vectors a_i for all $i \in \{1, 2, \ldots, p\}$ are distinct. Hence, the p vectors $a_i - a_1$ for all $i \in \{1, 2, \ldots, p\}$ are distinct as well (since subtracting a_1 from a bunch of vectors does not disturb their distinctness). Since all these p vectors belong to $\mathbf{A}'$, we thus conclude that the set $\mathbf{A}'$ has at least p elements. In other words, $|\mathbf{A}'| \geq p$. On the other hand, $\mathbf{A}'$ is an α-dimensional vector space (since $\dim(\mathbf{A}') = \alpha$). Hence, Proposition 10.4.4 (applied to α and $\mathbf{A}'$ instead of n and V) yields $|\mathbf{A}'| = 2^\alpha$. Thus, $2^\alpha = |\mathbf{A}'| \geq p$, so that $p \leq 2^\alpha$. Similarly, $q \leq 2^\beta$.

We have now found two integers $\alpha \in \mathbb{N}$ and $\beta \in \mathbb{N}$ such that $\alpha + \beta \leq n - 1$ and $p \leq 2^\alpha$ and $q \leq 2^\beta$. Thus, two such integers exist. This solves Exercise 10.5.1 **(a)**.

(b) Multiplying the equalities $p \leq 2^\alpha$ and $q \leq 2^\beta$, we obtain $pq \leq 2^\alpha \cdot 2^\beta = 2^{\alpha+\beta} \leq 2^{n-1}$ (since $\alpha + \beta \leq n - 1$). This solves Exercise 10.5.1 **(b)**. □

---

span $\{u_1, u_2, \ldots, u_k, a_1\}$ (since span $\{u_1, u_2, \ldots, u_k, a_1\}$ is a subspace of $\mathbb{F}_2^n$ and thus is closed under linear combination). In other words,

$$\operatorname{span}\{a_1, a_2, \ldots, a_p\} \subseteq \operatorname{span}\{u_1, u_2, \ldots, u_k, a_1\}.$$

Since $\mathbf{A} = \operatorname{span}\{a_1, a_2, \ldots, a_p\}$, we can rewrite this as

$$\mathbf{A} \subseteq \operatorname{span}\{u_1, u_2, \ldots, u_k, a_1\}. \tag{32}$$

On the other hand, the vectors $u_1, u_2, \ldots, u_k$ belong to $\mathbf{A}'$ and thus to $\mathbf{A}$ (since $\mathbf{A}' \subseteq \mathbf{A}$). The vector a_1 belongs to $\mathbf{A}$ as well. Hence, all $k+1$ vectors $u_1, u_2, \ldots, u_k, a_1$ belong to $\mathbf{A}$. Their linear combinations must therefore belong to $\mathbf{A}$ as well (since $\mathbf{A}$ is a subspace of $\mathbb{F}_2^n$ and thus is closed under linear combination). In other words,

$$\operatorname{span}\{u_1, u_2, \ldots, u_k, a_1\} \subseteq \mathbf{A}.$$

Combining this with (32), we find $\mathbf{A} = \operatorname{span}\{u_1, u_2, \ldots, u_k, a_1\}$. In other words, the $k+1$ vectors $u_1, u_2, \ldots, u_k, a_1$ span the vector space $\mathbf{A}$.

Using Exercise 10.5.1, we can now easily solve the remaining two of the four Oddtown/Eventown theorems:

Exercise 10.5.2. Let $n, m \in \mathbb{N}$. In a town with n inhabitants, there are m clubs, each of which has an odd number of members. Any two distinct clubs share an odd number of common members. No two distinct clubs have the exact same set of members. Prove that $m \leq 2^{\lfloor (n-1)/2 \rfloor}$.

Solution idea. Let S be the set of all the n inhabitants. Let $A_1, A_2, \ldots, A_m$ be the m clubs (regarded as subsets of S). Let $B_1, B_2, \ldots, B_m$ be the same m clubs, listed once again. Then, the m subsets $A_1, A_2, \ldots, A_m$ are distinct (since no two distinct clubs have the exact same set of members). Similarly, the m subsets $B_1, B_2, \ldots, B_m$ are distinct.

Moreover, the number $\left|A_i \cap B_j\right|$ is odd for every $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2, \ldots, m\}$. (Indeed, either A_i and B_j are two different clubs, in which case this follows from the assumption "any two distinct clubs share an odd number of common members"; or A_i and B_j are the same club, in which case this follows from the assumption "each club has an odd number of members".)

Thus, Exercise 10.5.1 **(a)** (applied to $p = m$ and $q = m$) shows that there exists integers $\alpha \in \mathbb{N}$ and $\beta \in \mathbb{N}$ such that $\alpha + \beta \leq n - 1$ and $m \leq 2^\alpha$ and $m \leq 2^\beta$. Consider these α and β.

WLOG assume that $\alpha \leq \beta$ (since otherwise, we can achieve this by swapping α with β). Hence, $2\alpha = \alpha + \underbrace{\alpha}_{\leq \beta} \leq \alpha + \beta \leq n - 1$, so that $\alpha \leq (n-1)/2$. Since α is an integer, this entails $\alpha \leq \lfloor (n-1)/2 \rfloor$. Thus, $m \leq 2^\alpha \leq 2^{\lfloor (n-1)/2 \rfloor}$ (since $\alpha \leq \lfloor (n-1)/2 \rfloor$). This solves Exercise 10.5.2. □

Exercise 10.5.3. Let $n, m \in \mathbb{N}$. In a town with n inhabitants, there are m clubs, each of which has an even number of members. Any two distinct clubs share an even number of common members. No two distinct clubs have the exact same set of members. Prove that $m \leq 2^{\lfloor n/2 \rfloor}$.

Solution idea. Invite an outsider to move into town and join every existing club. Thus,

- the town now has $n + 1$ inhabitants and m clubs;
- each club now has an odd number of members (since it had an even number of members before the outsider joined);
- any two distinct clubs now share an odd number of common members (since they shared an even number of common members before the outsider joined them both);
- no two distinct clubs have the exact same set of members (since this held before the outsider joined).

Thus, we can apply Exercise 10.5.2 to the new situation (keeping in mind that we must substitute $n+1$ for n). We obtain $m \leq 2^{\lfloor ((n+1)-1)/2 \rfloor} = 2^{\lfloor n/2 \rfloor}$ (since $(n+1)-1 = n$). This solves Exercise 10.5.3. $\square$

Remark 10.5.4. In both Exercise 10.5.2 and Exercise 10.5.3, the inequalities are sharp (i.e., equality is possible). Let me show how to obtain $m = 2^{\lfloor n/2 \rfloor}$ in Exercise 10.5.3:

Let the n inhabitants be $1, 2, \ldots, n$. Call a subset T of $\{1, 2, \ldots, n\}$ *blocky* if it has the property that an odd integer $i \in \{1, 2, \ldots, n\}$ belongs to T if and only if $i+1$ belongs to T. (For example, for $n = 5$, the blocky subsets are $\varnothing$, $\{1,2\}$, $\{3,4\}$ and $\{1,2,3,4\}$. The name "blocky" comes from the fact that the pairs $(1,2)$, $(3,4)$, $(5,6)$, ... are treated as "blocks", with each block either completely included in or completely excluded from the subset.)

There are $2^{\lfloor n/2 \rfloor}$ blocky subsets of $\{1, 2, \ldots, n\}$ (why?). If we use these blocky subsets as clubs, then the conditions of Exercise 10.5.3 are satisfied (why?), and we have $m = 2^{\lfloor n/2 \rfloor}$.

Constructing a similar example for equality in Exercise 10.5.2 is left as an exercise.

See [BabFra23] and [Stanle18, §13.2 and Exercises] for further results on and variants of Eventown/Oddtown as well as other combinatorial questions that can be resolved using linear algebra over finite fields.

Further applications of linear algebra can be found in [AndDos10, Chapter 12] and [AndDos12, Chapter 12].

10.6. Class problems

The following problems are to be discussed during class.

Exercise 10.6.1. Recall the Fibonacci sequence $(f_0, f_1, f_2, \ldots)$, defined recursively by

$$f_0 = 0, \qquad f_1 = 1, \qquad \text{and} \qquad f_n = f_{n-1} + f_{n-2} \text{ for all } n \geq 2.$$

We furthermore set $f_{-1} := 1$ (so that the equality $f_n = f_{n-1} + f_{n-2}$ also holds for $n = 1$).

Let A be the 2×2-matrix $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.

(a) Prove that $A^n = f_n A + f_{n-1} I_2$ for all $n \in \mathbb{N}$. (Here, I_2 denotes the 2×2 identity matrix.)

(b) Prove that $f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1}$ for all $n, m \in \mathbb{N}$. (This was Exercise 1.1.1 **(c)** on Worksheet 1.)

(c) Prove that $f_{n+1}f_{n-1} - f_n^2 = (-1)^n$ for each positive integer n. (This was Exercise 1.1.1 **(b)** on Worksheet 1.)

(d) Let v be the column vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Prove that $A^n v = \begin{pmatrix} f_n \\ f_{n+1} \end{pmatrix}$ for each $n \in \mathbb{N}$.

(e) Let $n, p \in \mathbb{N}$. Prove that

$$\sum_{k=0}^{n} \binom{n}{k} f_{k+p} = f_{p+2n}.$$

(f) Let $n, p \in \mathbb{N}$ be such that $p \geq n$. Prove that

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} f_{k+p} = (-1)^n f_{p-n}.$$

[**Hint:** For part **(e)**, take the equality $A^2 = A + I_2$ to the n-th power and expand the right hand side using the binomial formula.]

Exercise 10.6.2. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n$ be n numbers (or, more generally, elements of a field $\mathbb{F}$). Let $p_1, p_2, \ldots, p_n$ be n polynomials in $\mathbb{R}[X]$ (more generally, in $\mathbb{F}[X]$ where $\mathbb{F}$ is our field) with the property that

$$\deg(p_j) \leq j - 1 \qquad \text{for each } j \in \{1, 2, \ldots, n\}.$$

(In particular, p_1 is constant.)

(a) Prove that

$$\det\left(\left(p_j(a_i)\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}\right) = \left(\prod_{j=1}^{n} \left[X^{j-1}\right](p_j)\right) \cdot \det\left(\left(a_i^{j-1}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}\right).$$

(Recall that $\left[X^k\right] p$ means the X^k-coefficient of a polynomial p.)

(b) Prove that

$$\det\left(\left(a_i^{j-1}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}\right) = \prod_{1 \leq j < i \leq n} (a_i - a_j).$$

(This is known as the *Vandermonde determinant*.)

(c) Conclude that

$$\det\left(\left(p_j(a_i)\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}\right) = \left(\prod_{j=1}^{n} \left[X^{j-1}\right] p_j\right) \cdot \prod_{1 \leq j < i \leq n} (a_i - a_j).$$

(d) Prove that

$$\det\left(\left(\binom{a_i}{j-1}\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = \frac{\prod\limits_{1\leq j<i\leq n}\left(a_i-a_j\right)}{\prod\limits_{1\leq j<i\leq n}\left(i-j\right)}.$$

(Here, $\binom{a_i}{j-1}$ is a binomial coefficient, not a column vector.)

[**Hint:** Part **(a)** can be done in many ways, but the simplest is probably by factoring the matrix $\left(p_j\left(a_i\right)\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ as a product of two matrices. To prove part **(b)**, apply part **(a)** to n special polynomials $p_1, p_2, \ldots, p_n$, chosen strategically to make the matrix $\left(p_j\left(a_i\right)\right)_{1\leq i\leq n,\ 1\leq j\leq n}$ triangular. This is the nicest proof of the Vandermonde determinant, in my opinion, but there are many others that are not much worse.]

Exercise 10.6.3. Let $n \in \mathbb{N}$. Let $a_1, a_2, \ldots, a_n$ be n integers. Prove that

$$\prod_{1\leq i<j\leq n}\left(j-i\right) \mid \prod_{1\leq i<j\leq n}\left(a_j-a_i\right).$$

Exercise 10.6.4. You are given a row of 2023 lamps, each of which is either on or off. To "flip" a lamp means to change its state (i.e., turn it on if it was off and turn it off if it was on).

You have a tool that can flip 50 consecutive lamps on the row; but the tool can only be applied when the leftmost of these 50 lamps is on. (So you can choose any lamp that is on and has at least 49 further lamps to its right; then the tool turns it off and flips the next 49 lamps to its right.) You can use this tool as often as you want.

Prove the following:

(a) No matter how you use this tool, you will eventually run into a state where it can no longer be used (i.e., all lamps that have at least 49 lamps to their right are off).

(b) This final state does not depend on the choices you made (i.e., how you applied the tool), but only on the initial state.

Exercise 10.6.5. A complex number $z \in \mathbb{C}$ is said to be *algebraic* if there exists a nonzero polynomial $P \in \mathbb{Q}\left[X\right]$ (yes, a polynomial with rational coefficients) such that z is a root of P. (For instance, $\sqrt{2}$ and $\sqrt[3]{15}$ and all roots of $X^5 - 7X - 1$ are algebraic.)

Let u and v be two algebraic complex numbers. Prove that the numbers $u+v$ and uv are algebraic as well.

[**Hint:** Let $P \in \mathbb{Q}[X]$ and $Q \in \mathbb{Q}[X]$ be nonzero polynomials satisfying $P(u) = 0$ and $Q(v) = 0$, and let $n = \deg P$ and $m = \deg Q$ be their degrees. Now use linear algebra over $\mathbb{Q}$, treating $\mathbb{C}$ as an (infinite-dimensional) vector space over $\mathbb{Q}$. Let W be the vector subspace of $\mathbb{C}$ spanned by all the nm products of the form $u^i v^j$ with $i \in \{0,1,\ldots,n-1\}$ and $j \in \{0,1,\ldots,m-1\}$ over $\mathbb{Q}$ (that is, the elements of W are the $\mathbb{Q}$-linear combinations of these products). Prove that W is preserved under multiplication by u and by v (meaning that if $x \in W$, then $ux \in W$ and $vx \in W$). Conclude that the $nm+1$ powers

$$(u+v)^0, \ (u+v)^1, \ \ldots, \ (u+v)^{nm}$$

all belong to W. What now?]

10.7. Homework exercises

This homework set is optional and not to be graded.

Exercise 10.7.1. Let $n \in \mathbb{N}$. Let S be an n-element set. Let $A_1, A_2, \ldots, A_{n+2}$ be $n+2$ subsets of S. Prove that there exist two disjoint nonempty subsets I and J of $\{1,2,\ldots,n+2\}$ such that

$$\bigcup_{i \in I} A_i = \bigcup_{i \in J} A_i \qquad \text{and} \qquad \bigcap_{i \in I} A_i = \bigcap_{i \in J} A_i.$$

[**Hint:** Append an extra 1 at the end of each indicator vector.]

Exercise 10.7.2. Let $n > 1$ be an integer. Let S be an n-element set. Let d be a positive integer.

Let $A_1, A_2, \ldots, A_m$ be some proper subsets of S. Assume that every 2-element subset of S is contained (as a subset) in exactly d many of these subsets $A_1, A_2, \ldots, A_m$.

Prove that $m \geq n$.

Exercise 10.7.3. Let $a_1, a_2, \ldots, a_n$ be n numbers (or, more generally, elements of a field). Compute

$$\det\left(\left(a_{\min\{i,j\}}\right)_{1 \leq i \leq n,\ 1 \leq j \leq n}\right) = \det\begin{pmatrix} a_1 & a_1 & a_1 & \cdots & a_1 \\ a_1 & a_2 & a_2 & \cdots & a_2 \\ a_1 & a_2 & a_3 & \cdots & a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_n \end{pmatrix}.$$

Exercise 10.7.4. Let $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n$ be $2n$ reals.

(a) Compute

$$\det\left(\left(\alpha_i+\beta_j\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = \det\begin{pmatrix} \alpha_1+\beta_1 & \alpha_1+\beta_2 & \cdots & \alpha_1+\beta_n \\ \alpha_2+\beta_1 & \alpha_2+\beta_2 & \cdots & \alpha_2+\beta_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_n+\beta_1 & \alpha_n+\beta_2 & \cdots & \alpha_n+\beta_n \end{pmatrix}.$$

(b) Compute

$$\begin{aligned}&\det\left(\left(\sin\left(\alpha_i+\beta_j\right)\right)_{1\leq i\leq n,\ 1\leq j\leq n}\right) \\ &= \det\begin{pmatrix} \sin\left(\alpha_1+\beta_1\right) & \sin\left(\alpha_1+\beta_2\right) & \cdots & \sin\left(\alpha_1+\beta_n\right) \\ \sin\left(\alpha_2+\beta_1\right) & \sin\left(\alpha_2+\beta_2\right) & \cdots & \sin\left(\alpha_2+\beta_n\right) \\ \vdots & \vdots & \ddots & \vdots \\ \sin\left(\alpha_n+\beta_1\right) & \sin\left(\alpha_n+\beta_2\right) & \cdots & \sin\left(\alpha_n+\beta_n\right) \end{pmatrix}.\end{aligned}$$

[**Hint:** The cases $n \leq 2$ might be somewhat misleading...]

Exercise 10.7.5. Let $n \in \mathbb{N}$, and let a and b be two numbers. Prove that

$$\det\underbrace{\begin{pmatrix} b & -a & -a & \cdots & -a \\ a & b & -a & \cdots & -a \\ a & a & b & \cdots & -a \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a & a & a & \cdots & b \end{pmatrix}}_{n\times n\text{-matrix}} = \frac{\left(b+a\right)^n + \left(b-a\right)^n}{2}.$$

(Here, all entries below the diagonal are a's; all entries on the diagonal are b's; all entries above the diagonal are $-a$'s.)

Exercise 10.7.6. Let

$$A = \left(a_{i,j}\right)_{1\leq i\leq n,\ 1\leq j\leq n} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

be an $n \times n$-matrix with real entries. Assume that

$$a_{i,i} \geq \sum_{\substack{j\in\{1,2,\ldots,n\};\\ j\neq i}} \left|a_{i,j}\right| \qquad \text{for each } i \in \{1,2,\ldots,n\}.$$

(Such a matrix A is said to be *weakly diagonally dominant*.) Prove that

$$\det A \geq \prod_{i=1}^{n} \left(a_{i,i} - \sum_{j=i+1}^{n} \left| a_{i,j} \right| \right).$$

The following exercise is a generalization of Exercise 10.4.1:

Exercise 10.7.7. Let p be a prime number.

Let $n, m \in \mathbb{N}$. In a town with n inhabitants, there are m clubs, each of which has a number of members that is not divisible by p. Any two distinct clubs have a number of common members that is divisible by p. Prove that $m \leq n$.

We note that a similar generalization of Exercise 10.4.2 does not work, as (e.g.) the example of $p = 3$ and $n = 5$ and $m = \binom{5}{3} = 10$ shows (here, each three inhabitants form a club).

Exercise 10.7.8. An ornithological handbook classifies birds by 100 attributes; each bird either has a given attribute or does not have it. Two birds are said to be *dissimilar* if they differ in at least 51 attributes.

(a) Show that the handbook cannot contain 51 birds all dissimilar from each other.

(b) (harder:) Can it contain 50 such birds?

(Tournament of Towns 14.42)

[**Hint:** In part **(a)**, you can replace 100 and 51 by $2n$ and $n+1$, respectively, where n is an even integer. In part **(b)**, you can replace 50 by n, but you must additionally require that $n > 4$.]

Exercise 10.7.9. Let $n, m \in \mathbb{N}$. In a town with n inhabitants, there are m red clubs $R_1, R_2, \ldots, R_m$ and m blue clubs $B_1, B_2, \ldots, B_m$. Assume that $|R_i \cap B_i|$ is odd for each $i \in \{1, 2, \ldots, m\}$. Assume further that $\left|R_i \cap B_j\right|$ is even for each $i, j \in \{1, 2, \ldots, m\}$ satisfying $i < j$. Prove that $m \leq n$.

Exercise 10.7.10. Let $n, b, r \in \mathbb{N}$. In a town with n inhabitants, there are r red clubs and b blue clubs. Each red club has an even number of members, whereas each blue club has an odd number of members. Any two distinct clubs (no matter their color) share an even number of common members. No two distinct clubs have the exact same set of members. Prove that $r \leq 2^{\lfloor (n-b)/2 \rfloor}$.

References

[Aluffi21] Paolo Aluffi, *Algebra: Notes from the Underground*, Cambridge University Press 2021.

[AndDos10] Titu Andreescu, Gabriel Dospinescu, *Problems from the Book*, 2nd edition, XYZ Press 2010.

[AndDos12] Titu Andreescu, Gabriel Dospinescu, *Straight from the Book*, XYZ Press 2012.

[Axler23] Sheldon Axler, *Linear Algebra Done Right*, 4th edition, Springer 2023.

[BabFra23] László Babai, Péter Frankl, *Linear algebra methods in combinatorics*, 7 January 2023.
`https://people.cs.uchicago.edu/~laci/babai-frankl-book2022.pdf`

[Bollob10] Béla Bollobás, *The Art of Mathematics: Coffee Time in Memphis*, Cambridge University Press 2010.

[Bollob22] Béla Bollobás, *The Art of Mathematics – Take Two: Tea Time in Cambridge*, Cambridge University Press 2022.

[Camero08] Peter J. Cameron, *Notes on Linear Algebra*, 2008.
`https://cameroncounts.files.wordpress.com/2013/11/linalg.pdf`

[GelAnd17] Răzvan Gelca, Titu Andreescu, *Putnam and Beyond*, 2nd edition, Springer 2017.

[Griffin20] Christopher Griffin, *Intermediate Linear Algebra*, Version 2.1, 13 December 2020.
`https://sites.psu.edu/griffinch/lecture_notes/`

[Grinbe15] Darij Grinberg, *Notes on the combinatorial fundamentals of algebra*, 15 September 2022, arXiv:2008.09862v3.

[Grinbe18] Darij Grinberg, *Math 4707 Spring 2018 (Darij Grinberg): homework set 4 with solutions*.
`https://www.cip.ifi.lmu.de/~grinberg/t/18s/hw4.pdf`

[Grinbe20a] Darij Grinberg, *Math 235: Mathematical Problem Solving*, 10 August 2021.
`https://www.cip.ifi.lmu.de/~grinberg/t/20f/mps.pdf`

[Grinbe20b] Darij Grinberg, *A note on bilinear forms*, July 9, 2020.
`https://www.cip.ifi.lmu.de/~grinberg/algebra/bilf.pdf`

[Grinbe21] Darij Grinberg, *An Introduction to Algebraic Combinatorics [Math 701, Spring 2021 lecture notes]*, 19 December 2022.
https://www.cip.ifi.lmu.de/~grinberg/t/21s/lecs.pdf

[Grinbe23a] Darij Grinberg, *An introduction to the algebra of rings and fields (Text for Math 332 Winter 2023 at Drexel University)*, 27 August 2023.
https://www.cip.ifi.lmu.de/~grinberg/t/23wa/23wa.pdf

[Grinbe23b] Darij Grinberg, *An introduction to graph theory (Text for Math 530 in Spring 2022 at Drexel University)*, arXiv:2308.04512v1.

[Griffi20] Christopher Griffin, *Intermediate Linear Algebra*, Version 2.1, 13 December 2020.
https://sites.psu.edu/griffinch/lecture_notes/

[Kratte99] Christian Krattenthaler, *Advanced Determinant Calculus*, Séminaire Lotharingien Combin. 42 (1999) (The Andrews Festschrift), paper B42q, 67 pp., arXiv:math/9902004v3.

[Kuttle22] Kenneth Kuttler, *Elementary Linear Algebra*, November 15, 2023.
https://klkuttler.com/

[LaNaSc16] Isaiah Lankham, Bruno Nachtergaele, Anne Schilling, *Linear Algebra As an Introduction to Abstract Mathematics*, 2016.
https://www.math.ucdavis.edu/~anne/linear_algebra/mat67_course_notes.pdf

[Prasol94] Viktor V. Prasolov, *Problems and Theorems in Linear Algebra*, Translations of Mathematical Monographs, vol. #134, AMS 1994.

[Stanle18] Richard P. Stanley, *Algebraic Combinatorics: Walks, Trees, Tableaux, and More*, 2nd edition, Springer 2018.
See http://www-math.mit.edu/~rstan/algcomb/index.html for errata.

[Steinb06] Mark Steinberger, *Algebra*, 31 August 2006.
https://math.hawaii.edu/~tom/algebra.pdf

[StoLui18] Michael Stoll, Ronald van Luijk, *Linear Algebra I*, 2 September 2018.
https://websites.math.leidenuniv.nl/algebra/linalg1.pdf

[Strick21] Neil Strickland, *Linear Mathematics for Applications*, 11 February 2020.
https://neilstrickland.github.io/linear_maths/

[Treil21] Serge Treil, *Linear Algebra Done Wrong*, 11 January 2021.
https://sites.google.com/a/brown.edu/sergei-treil-homepage/linear-algebra-done-wrong

[Zhao09] Yufei Zhao, *Determinants: Evaluation and Manipulation*, September 22, 2009.
`https://yufeizhao.com/olympiad/det_eval_man.pdf`